

# SEARCH REQUEST FORM

7.511

Requestor's Name: Nancy Johnson Serial Number: 091047,652  
Date: 7-20-98 Phone: 305-5860 Art Unit: 1642

## Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

Please search SEQ ID NO:1-3

Please do an interference search of

Seq ID NO:1-3.

Return on paper print out.

Thanks!

Nancy J.

1998 JUL 20 PM 3:20

## STAFF USE ONLY

Date completed: 7/22/98  
Searcher: S. L. Grand  
Terminal time: \_\_\_\_\_  
Elapsed time: \_\_\_\_\_  
CPU time: \_\_\_\_\_  
Total time: \_\_\_\_\_  
Number of Searches: \_\_\_\_\_  
Number of Databases: \_\_\_\_\_

Search Site  
\_\_\_\_\_ STIC  
\_\_\_\_\_ CM-1  
\_\_\_\_\_ Pre-S  
Type of Search  
\_\_\_\_\_ N.A. Sequence  
\_\_\_\_\_ A.A. Sequence  
\_\_\_\_\_ Structure  
\_\_\_\_\_ Bibliographic

Vendors  
\_\_\_\_\_ IG Suite  
\_\_\_\_\_ STN  
\_\_\_\_\_ Dialog  
\_\_\_\_\_ APS  
\_\_\_\_\_ Geninfo  
\_\_\_\_\_ SDC  
\_\_\_\_\_ DARC/Questel  
\_\_\_\_\_ Other

**THIS PAGE BLANK (USPTO)**

\*\*\*\*\*  
MREH (TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm  
Run on: Tue Jul 21 14:53:17 1998; MasPar time 828.83 Seconds  
Tabular output not generated. 1312.667 Million cell updates/sec

Title: >US-09-047-652A-1  
Description: (1-652) from US09047652A.seq  
Perfect Score: 652  
N.A. Sequence: 1 CCACGCCGAAGGTCTCCGCT.....GTTCTTGGACATGGAATTT 652  
Comp: GGTGCCGCTCCAGAGCGGA.....CAAGAACCTGTACCTTAA

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0  
Searched: 457396 seqs, 834342348 bases x 2  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: embl54  
1:em\_ba 2:em\_htg 3:em\_hum1 4:em\_hum2 5:em\_in 6:em\_om  
7:em\_or 8:em\_ov 9:em\_pat 10:em\_pl 11:em\_ro 12:em\_vi  
genbank106  
Database: 13:gb\_ba 14:gb\_htg 15:gb\_in 16:gb\_om 17:gb\_ov 18:gb\_pat  
19:gb\_ph 20:gb\_pl 21:gb\_pr1 22:gb\_pr2 23:gb\_ro 24:gb\_st  
25:gb\_sy 26:gb\_un 27:gb\_vi

Statistics: Mean 10.377; Variance 5.497; scale 1.888  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match	Length	DB ID	Description
1	643	98.6	821	21	HUMHPBS Human peripheral benzo
2	542	83.1	639	21	HUMBENZA Human peripheral benzo
3	403	61.8	684	21	HSPBR4 Human peripheral benzo
4	403	61.8	160559	14	HS1191B2 Human DNA sequence ***
5	402	61.7	4258	21	HSU12421 Human mitochondrial be
6	384	58.9	152843	14	HS526114 Human DNA sequence ***
7	288	44.2	821	16	BOVPBRIBP Calf peripheral-type b
8	282	43.3	626	23	MUSMPBR Mouse peripheral-type b
9	276	42.3	781	23	RATPKBSX Rat peripheral-type be
10	274	42.0	856	23	MUSPTBR Mouse mRNA for periphe
11	128	19.6	645	21	HSPBR3 Human peripheral benzo
12	115	17.6	3434	23	RATPTBZR02 Rat peripheral-type be
13	106	16.3	487	21	HSPBR2 Human peripheral benzo
14	39	6.0	7218	18	I66494 Sequence 14 from paten
15	37	5.7	1084	23	CGU12420 Cricetulus griseus mit

C	16	5.5	215	18	128278	Sequence 5 from patent	1.22e-06
	17	5.5	1525	23	MMU12419	Mus musculus mitochond	1.22e-06
	18	5.4	215	18	128278	Sequence 5 from patent	4.82e-06
	19	4.9	3084	13	S71770	carotenoid gene cluste	2.72e-04
	20	4.9	9298	13	SPCRT	R.sphaeroides crt gene	2.72e-04
	21	3.8	565	18	E04076	gDNA encoding envelope	1.71e+00
C	22	3.8	1284	22	HSF000972	Homo sapiens calcium-a	1.71e+00
C	23	3.8	1284	22	AF022150	Homo sapiens intermedi	1.71e+00
C	24	3.8	1328	13	AE0AMOA	A.hydrophila amonabact	1.71e+00
	25	3.8	1684	21	HUMTTP	Human tristetraproline	1.71e+00
	26	3.8	1746	21	HUMGOS24A	H.sapiens zinc finger	1.71e+00
C	27	3.8	1982	22	AF033021	Homo sapiens intermed	1.71e+00
C	28	3.8	2238	22	AF022797	Homo sapiens intermed	1.71e+00
	29	3.8	3889	22	HUMGOS24B	Homo sapiens zinc fing	1.71e+00
	30	3.7	196	13	MSU77950	Mycobacterium smegmati	5.42e+00
	31	3.7	1120	13	AF005905	Mycobacterium smegmati	5.42e+00
	32	3.7	1804	22	HOSAWSCR1	Homo sapiens WSCR4 gen	5.42e+00
C	33	3.7	2926	13	S80432	ntfB, ntrC [Azospirill	5.42e+00
	34	3.7	4943	22	AB006629	Homo sapiens mRNA for	5.42e+00
	35	3.7	5350	23	RNU27319	Rattus norvegicus type	5.42e+00
	36	3.7	9483	23	MUSOPS	Mouse opsin (MOPS) gen	5.42e+00
C	37	3.7	134914	22	AC004542	Homo sapiens PAC clone	5.42e+00
C	38	3.5	69	18	I41362	Sequence 142 from pate	1.67e+01
C	39	3.5	1697	21	HUMALPPC1	Human placental alkali	1.67e+01
C	40	3.5	2487	21	HUMALPIGC	Human alkaline phospho	1.67e+01
C	41	3.5	2688	21	HUMALPPP	Human placental alkali	1.67e+01
C	42	3.5	5115	25	CVU89938	Cloning vector pSEAP2-	1.67e+01
	43	3.5	10068	13	RCCRTAK	Rhodobacter capsulatus	1.67e+01
	44	3.5	119307	22	AC004134	Homo sapiens chromosom	1.67e+01
C	45	3.5	276261	14	HSAC000406	*** SEQUENCING IN PROG	1.67e+01

ALIGNMENTS

RESULT 1 HUMHPBS 821 bp mRNA PRI 20-DEC-1993  
LOCUS Human peripheral benzodiazepine receptor (hpbs) mRNA, complete cds.  
DEFINITION M36035  
ACCESSION  
NID g184333  
KEYWORDS peripheral benzodiazepine receptor.  
SOURCE Human cDNA to mRNA.  
ORGANISM Homo sapiens  
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 821)  
AUTHORS Riond, J., Mattel, M.G., Kaghad, M., Dumont, X., Guillemot, J.C., Le Fur, G., Caput, D. and Ferrara, P.  
TITLE Molecular cloning and chromosomal localization of a human peripheral-type benzodiazepine receptor  
JOURNAL Eur. J. Biochem. 195 (2), 305-311 (1991)  
MEDLINE 91146565  
REFERENCE 2 (bases 1 to 821)  
AUTHORS Riond, J.  
TITLE Direct Submission  
JOURNAL Submitted (27-JUN-1990) J. Riond, Sanofi Elf Bio-Recherches, BP137, 31328 Labège Cedex, France  
FEATURES Location/Qualifiers  
source  
1..821  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_line="hystiocytic lymphoma monocyte-like cell line U937"  
/clone="p-hpBS11"  
<1..811  
/note="peripheral benzodiazepine receptor mRNA"  
62..571  
/note="peripheral benzodiazepine receptor"  
/codon\_start=1  
/db\_xref="PID:g306883"  
/translation="MAPPPVPMAGFTLAPSLGCFVGSREVFHGEGLRWYAGLQKPSWHP  
PHWVLPVWGTLYSAMGYSLVMKELGGFTKAVPLGLYGLQALNWAWPPIFFGA  
RQMGWALVDLLLVSGAAATTVAVYQVSPLAARLLYPYLAWLAFATTLNVCVWRDNG  
WHGGRRLPE"

polyA_signal 800..805									
BASE COUNT		118 a	271 c	260 g	171 t	1	others		
ORIGIN		Chromosome 22, map position q13.3.							
Query Match		98.6%; Score 643; DB 21; Length 821;							
Best Local Similarity		99.2%; Pred. No. 0.00e+00;							
Matches		647; Conservative 0; Mismatches 5; Indels 0; Gaps 0;							
Db	139	CCACGGCGAGGGTCTCCGCTGGTACGCCGGCCCTGCAGAGCCCTCGTGGCACCCGCCCA	198						
QY	1	CCACGGCGAAGGCTCTCCGCTGGTACGCCGGCCCTGCAGAGCCCTCGTGGCACCCGCCCA	60						
Db	199	CTGGGTGCTGGGCCCTGTCTGGGGCACGCTCTACTCAGCCATGGGGTACGGCTCCTACCT	258						
QY	61	CTGGGTGCTGGGCCCTGTCTGGGGCACGCTCTACTCAGCCATGGGGTACGGCTCCTACCT	120						
Db	259	GGTCTGGAAGAGCTGGGAGGCTTCACAGAGAGGCTGTGGTTCCCTGGGCTCTACAC	318						
QY	121	GGTCTGGAAGAGCTGGGAGGCTTCACAGAGAGGCTGTGGTTCCCTGGGCTCTACAC	180						
Db	319	TGGCAGCTGGCCCTGAACCTGGGCATGGCCCCCCCATCTTCTTTGGTGCCGACAAATGGG	378						
QY	181	TGGCAGCTGGCCCTGAACCTGGGCATGGCCCCCCCATCTTCTTTGGTGCCGACAAATGGG	240						
Db	379	CTGGGCTTGGTGATCTCCTGCTGGTCAGTGGGGCGGCGNCGCCACTACCGTGGCCTG	438						
QY	241	CTGGGCTTGGTGATCTCCTGCTGGTCAGTGGGGCGGCGGACCACTACCGTGGCCTG	300						
Db	439	GTACCAGGTGAGCCCGCTGCCCGCCCGCTGCTCTACCCCTACCTGGCCTGGCCTT	498						
QY	301	GTACCAGGTGAGCCCGCTGCCCGCCCGCTGCTCTACCCCTACCTGGCCTGGCCTT	360						
Db	499	CGGACCACACTCAACTACTGCGTATGGCGGGACACCATGGCTGGCATGGGGACGGCG	558						
QY	361	CACGACCACACTCAACTACTGCGTATGGCGGGACACCATGGCTGGCATGGGGACGGCG	420						
Db	559	GCTGCCAGAGTGAGTGCCCGCCCCACCAGGAGCTGCAGCTGCACCAGCAGGTGCCATCAC	618						
QY	421	GCTGCCAGAGTGAGTGCCCGCCCCACCAGGAGCTGCAGCTGCACCAGCAGGTGCCATCAC	480						
Db	619	GCTTGATGTGGTGCCGCTCACGCTTTTCATGACCACTGGGCTGCTAGTCTGTACAGGC	678						
QY	481	GCTTGATGTGGTGCCGCTCACGCTTTTCATGACCACTGGGCTGCTAGTCTGTACAGGC	540						
Db	679	CTTGGCCAGGGGTACAGAGCTTCAGAGGTTGCCACCTGAGCCCCACCGGGAGC	738						
QY	541	CTTGGCCAGGGGTACAGAGCTTCAGAGGTTGCCACCTGAGCCCCACCGGGAGC	600						
Db	739	AGTGCTGTGCTTTCTGTCATGCTTAGAGCATGTTCTTGGAAATTT	790						
QY	601	AGTGCTGTGCTTTCTGTCATGCTTAGAGCATGTTCTTGGAAATTT	652						
RESULT 2									
LOCUS		HUMBENZA	639 bp	mrna	PRI	20-MAY-1994			
DEFINITION		Human peripheral benzodiazepine receptor related mRNA sequence.							
ACCESSION		L21950							
NID		g483401							
KEYWORDS		benzodiazepine receptor; peripheral benzodiazepine receptor.							
SOURCE		Homo sapiens cDNA to mRNA.							
ORGANISM		Homo sapiens							
REFERENCE		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homidae; Homo.							
AUTHORS		1 (bases 1 to 639)							
TITLE		Lin,D., Chang,Y.J., Strauss,J.F. and Miller,W.L.							
		The human peripheral benzodiazepine receptor gene: cloning and							
		characterization of alternative splicing in normal tissues and in a							
		patient with congenital lipoid adrenal hyperplasia							
JOURNAL		Genomics 18 (3), 643-650 (1993)							
MEDLINE		94140364							
FEATURES		Location/Qualifiers							
source		1..639							
		/organism="Homo sapiens"							

/db_xref="taxon:9606"									
CDS		161..469							
		/codon_start=1							
		/product="peripheral benzodiazepine receptor related protein"							
		/db_xref="PID:g488420"							
		/translation="MAPHLLWCTNGLGGSPAGQWGGSHYRGLVPGEAGRPALPLPGLAGLRDHTQLLRMAGQPLAWGTAAARVSARPTRDCSCTSRCHHACDVAVTLS"							
		"							
polyA_site		639	200 g		138 t				
BASE COUNT		100 a	201 c	200 g	138 t				
ORIGIN									
Query Match		83.1%; Score 542; DB 21; Length 639;							
Best Local Similarity		99.5%; Pred. No. 0.00e+00;							
Matches		545; Conservative 0; Mismatches 3; Indels 0; Gaps 0;							
Db	61	GGTACGGCTCCTACCTGGTCTGGAAAGAGCTGGGAGGCTTCACAGAGAGGCTGTGGTTC	120						
QY	105	GGTACGGCTCCTACCTGGTCTGGAAAGAGCTGGGAGGCTTCACAGAGAGGCTGTGGTTC	164						
Db	121	CCCTGGGCCTCTACACTGGGCAGCTGGCCCTGAACCTGGGCATGGCCCCCATCTTCTTTG	180						
QY	165	CCCTGGGCCTCTACACTGGGCAGCTGGCCCTGAACCTGGGCATGGCCCCCATCTTCTTTG	224						
Db	181	GTGCCCCGACAAATGGGCTGGGCTTGGTGATCTCCTGCTGGTCAGTGGGGCGGGCAG	240						
QY	225	GTGCCCCGACAAATGGGCTGGGCTTGGTGATCTCCTGCTGGTCAGTGGGGCGGGCAG	284						
Db	241	CCACTACCGTGGCCTGGTACCAGGTGAGCCCGCTGGCCGCCCTGCTCTACCCCTACC	300						
QY	285	CCACTACCGTGGCCTGGTACCAGGTGAGCCCGCTGGCCGCCCTGCTCTACCCCTACC	344						
Db	301	TGGCCTGGCTGGCCTTCGCGACCACACTCAACTACTGCGTATGGCGGGACAACCATGGCT	360						
QY	345	TGGCCTGGCTGGCCTTCACGACCACACTCAACTACTGCGTATGGCGGGACAACCATGGCT	404						
Db	361	GGCATGGGGGACGGCGGTGCCAGAGTGAGTGCCCGGCCACCAGGGACTGCAGCTGCAC	420						
QY	405	GGCGTGGGGGACGGCGGTGCCAGAGTGAGTGCCCGGCCACCAGGGACTGCAGCTGCAC	464						
Db	421	CAGCAGGTGCCATCACGCTTGTGATGGTGGCCGTGCAGCTTTCATGACCACTGGGCT	480						
QY	465	CAGCAGGTGCCATCACGCTTGTGATGGTGGCCGTGCAGCTTTCATGACCACTGGGCT	524						
Db	481	GCTAGTCTGTACAGGGCCTTGCCCCAGGGGTACAGAGCTTCAGAGGTTGCCCCACCTGA	540						
QY	525	GCTAGTCTGTACAGGGCCTTGCCCCAGGGGTACAGAGCTTCAGAGGTTGCCCCACCTGA	584						
Db	541	GCCCCACCCGGGAGCAGTGCTGCTTTCTGTCATGCTTAGAGCATGTTCTTGGGAACA	600						
QY	585	GCCCCACCCGGGAGCAGTGCTGCTTTCTGTCATGCTTAGAGCATGTTCTTGGGAACA	644						
Db	601	TGGAATTT	608						
QY	645	TGGAATTT	652						
RESULT 3									
LOCUS		HSPBR4	684 bp	DNA	PRI	20-MAY-1994			
DEFINITION		Human peripheral benzodiazepine receptor gene, exon 4.							
ACCESSION		L21954							
NID		g483405							
KEYWORDS		benzodiazepine receptor; peripheral benzodiazepine receptor.							
SEGMENT		4 of 4							
SOURCE		Homo sapiens DNA.							
ORGANISM		Homo sapiens							
REFERENCE		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homidae; Homo.							
AUTHORS		1 (bases 1 to 684)							
TITLE		Lin,D., Chang,Y.J., Strauss,J.F. and Miller,W.L.							
		The human peripheral benzodiazepine receptor gene: cloning and							





Db 71540 CCAGGCCAGGTAGGGGTAGAGCAGCGGGCGGCCAGCGGGCTCACCTGGTACCAGGCCAC 71599  
|||||  
Cp 352 CCAGGCCAGGTAGGGGTAGAGCAGCGGGCGGCCAGCGGGCTCACCTGGTACCAGGCCAC 293  
|||||  
Db 71600 GGTAGTGGTGCCTGCCCGCCCACTGACCAGCAGGAGATCCACCAAGGCC 71648  
|||||  
Cp 292 GGTAGTGGTGCCTGCCCGCCCACTGACCAGCAGGAGATCCACCAAGGCC 244  
|||||  
RESULT 5  
LOCUS HSU12421 4258 bp DNA PRI 14-DEC-1995  
DEFINITION Human mitochondrial benzodiazepine receptor (MBR) gene, complete cds.  
ACCESSION U12421  
NID 9529945  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 2080 to 3683)  
AUTHORS Yakovlev,A.G., Ruffo,M., Jurka,J. and Krueger,K.E.  
TITLE Comparison of repetitive elements in the third intron of human and rodent mitochondrial benzodiazepine receptor-encoding genes  
JOURNAL Gene 155 (2), 201-205 (1995)  
MEDLINE 95237610  
REFERENCE 2 (bases 1 to 4258)  
AUTHORS Krueger,K.E.  
TITLE Direct Submission  
JOURNAL Submitted (19-JUL-1994) Karl E. Krueger, Dept. of Cell Biology, Georgetown University School of Medicine, Washington, DC 20007, USA  
FEATURES  
source 1..4258  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="22q13.3"  
intron <1..99  
/number=1  
/evidence=experimental  
exon 100..310  
/number=2  
/evidence=experimental  
/product="mitochondrial benzodiazepine receptor"  
100..128  
/db\_xref="PID:9529946"  
129..3872  
/gene="MBR"  
CDS join(129..310,1941..2079,3684..3872)  
/gene="MBR"  
/codon\_start=1  
/evidence=experimental  
/product="mitochondrial benzodiazepine receptor"  
/db\_xref="PID:9529946"  
/translation="MAPWPVPAMGFTLAPSLGCFVGSRFVHGEGLRWYAGLQKPSWHP  
PHWVLGPVWGTLYSAMGYGSYLVMKELGGFTEKAVPLGLYTGQLALNWAWPPIFFGA  
RQMGWALVDLLVSGAAATTVAWYQVSPLAARLLYPYLAWLAFTTLNYCVRDNHG  
WRGRRRLPE"  
intron 311..1940  
/gene="MBR"  
/number=2  
/evidence=experimental  
exon 1941..2079  
/gene="MBR"  
/number=3  
/evidence=experimental  
intron /product="mitochondrial benzodiazepine receptor"  
2080..3683  
/gene="MBR"  
/number=3  
/evidence=experimental  
repeat\_region 2401..3306  
/note="4 Alus"  
/rpt\_family="Alu"  
repeat\_region 3497..3581

exon /rpt\_family="MIR"  
3684..4122  
/number=4  
/evidence=experimental  
/product="mitochondrial benzodiazepine receptor"  
3873..4122  
BASE COUNT 794 a 1246 c 1221 g 997 t  
ORIGIN  
Query Match 61.7%; Score 402; DB 21; Length 4258;  
Best Local Similarity 99.8%; Pred. No. 2.43e-287;  
Matches 408; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
Db 3683 GGCCTTGGTGGATCTCCTGCTGGTCACTGGGGCGGGCAGCCACTACCGTGGCCTGGTA 3742  
|||||  
QY 244 GGCCTTGGTGGATCTCCTGCTGGTCACTGGGGCGGGCAGCCACTACCGTGGCCTGGTA 303  
|||||  
Db 3743 CCAGGTGAGCCCGCTGGCCGCCCGCTCTACCCCTACCTGGCTGGCTTTCAC 3802  
|||||  
QY 304 CCAGGTGAGCCCGCTGGCCGCCCGCTCTACCCCTACCTGGCTGGCTTTCAC 363  
|||||  
Db 3803 GACCACACTCAACTACTGCGTATGGGGGACAACCATGGCTGGGGGACGGCGGCT 3862  
|||||  
QY 364 GACCACACTCAACTACTGCGTATGGGGGACAACCATGGCTGGGGGACGGCGGCT 423  
|||||  
Db 3863 GCCAGAGTGAGTCCCGGGCCACCAGGAGTGCAGCTGCACCAGCAGGTGCCATCAGCT 3922  
|||||  
QY 424 GCCAGAGTGAGTCCCGGGCCACCAGGAGTGCAGCTGCACCAGCAGGTGCCATCAGCT 483  
|||||  
Db 3923 TGTGATGTGGTGGCGGTCACGCTTTTCATGACCACCTGGGGCTGCTAGTCTGTACAGGCGCTT 3982  
|||||  
QY 484 TGTGATGTGGTGGCGGTCACGCTTTTCATGACCACCTGGGGCTGCTAGTCTGTACAGGCGCTT 543  
|||||  
Db 3983 GGCCAGGGGTACAGAGAGCTTTCAGAGTGGCCCCCAC-TGAGCCCCACCCGGGAGCAGT 4041  
|||||  
QY 544 GGCCAGGGGTACAGAGAGCTTTCAGAGTGGCCCCCACCTGAGCCCCACCCGGGAGCAGT 603  
|||||  
Db 4042 GTCCTGTGCTTCTGTCATGCTTAGAGCATGTTCTTGAACATGGAATTT 4090  
|||||  
QY 604 GTCCTGTGCTTCTGTCATGCTTAGAGCATGTTCTTGAACATGGAATTT 652  
|||||  
RESULT 6  
LOCUS HS526I14 152843 bp DNA HTG 13-JAN-1997  
DEFINITION Human DNA sequence \*\*\* SEQUENCING IN PROGRESS \*\*\* from clone 526I14; HTGS phase 1.  
ACCESSION 282214  
NID 91666406  
KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 152843)  
AUTHORS Buck,D.  
TITLE Direct Submission  
JOURNAL Submitted (06-NOV-1996) Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and  
the release of this data is based on the understanding that the sequence may change as work continues. The sequence may be contaminated  
with foreign sequence from E.coli, yeast, vector, phage etc. Order of segments is not known; 800 n's separate segments. Unfinished sequence: dJ526I14 Contig\_ID: 00003 Length: 15560 bp Unfinished sequence: dJ526I14 Contig\_ID: 00229 Length: 791 bp Unfinished sequence: dJ526I14 Contig\_ID: 00240 Length: 810 bp Unfinished sequence: dJ526I14 Contig\_ID: 00248 Length: 750 bp Unfinished sequence: dJ526I14 Contig\_ID: 00250 Length: 763 bp Unfinished sequence: dJ526I14 Contig\_ID: 00263 Length: 794 bp Unfinished

sequence: dJ526I14 Contig\_ID: 00278 Length: 1221 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00281 Length: 1174 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00287 Length: 1142 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00295 Length: 1145 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00308 Length: 826 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00327 Length: 1117 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00342 Length: 1668 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00351 Length: 947 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00359 Length: 1163 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00369 Length: 1092 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00393 Length: 1262 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00408 Length: 2766 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00423 Length: 1714 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00435 Length: 2050 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00453 Length: 2436 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00547 Length: 3304 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00574 Length: 4133 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00610 Length: 4473 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00647 Length: 4540 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00680 Length: 3551 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00752 Length: 3422 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00793 Length: 4738 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00831 Length: 4862 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00870 Length: 2306 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00944 Length: 7105 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00995 Length: 6767 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 01058 Length: 4554 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 01131 Length: 8507 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 01204 Length: 9125 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00899 Length: 10714 bp Unfinished  
sequence: dJ526I14 Contig\_ID: 00896 Length: 751 bp.

\*\*\* WARNING: Phase 1 High Throughput Genome Sequence \*\*\*

\*\*\* This sequence is unfinished. When sequencing is complete,  
\* the sequence data presented in this record will be replaced  
\*by a single finished sequence with the same accession number.

FEATURES

source  
1. 152843  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="526I14"  
/chromosome="22"

BASE COUNT 28005 a 34564 c 33570 g 27611 t 29093 others

ORIGIN

Query Match 58.9%; Score 384; DB 14; Length 152843;  
Best Local Similarity 98.1%; Pred. No. 5.93e-273;  
Matches 404; Conservative 0; Mismatches 5; Indels 3; Gaps 3;

Db 134121 GGCCTTGGTGATCTCCTGCTGGTNCANGTGGGGCGCGNNGCCACTACCGTGGCCTGG 134180  
|||||  
QY 244 GGCCTTGGTGATCTCCTGCTGGT-CA-GTGGGGCGCGCGCAGCCACTACCGTGGCCTGG 301

Db 134181 TACCAGGTGAGCCCGCTGGCCGCCCGCTGCTATACCCCTACCTGGCCTGGCTGGCCTTC 134240  
|||||  
QY 302 TACCAGGTGAGCCCGCTGGCCGCCCGCTGCTCTACCCCTACCTGGCCTGGCTGGCCTTC 361

Db 134241 ACGACCACACTCAACTACTCGGTATGGCGGGACAAACATGGCTGGCGTGGGGACGGCGG 134300  
|||||  
QY 362 ACGACCACACTCAACTACTCGGTATGGCGGGACAAACATGGCTGGCGTGGGGACGGCGG 421

Db 134301 CTGCCAGAGTGAGTGCCCGGCCCCACCAGGGAGCTGCAGCTGCACCAGCAGGTGCCATCAGG 134360  
|||||  
QY 422 CTGCCAGAGTGAGTGCCCGGCCCCACCAGGGAGCTGCAGCTGCACCAGCAGGTGCCATCAGG 481

Db 134361 CTTGNTGATGTGGTGGCGCTACGCTTTTCATGACCAGTGGGCCTGCTAGTCTGTCAGGGC 134420  
|||||  
QY 482 CTTG-TGATGTGGTGGCGCTACGCTTTTCATGACCAGTGGGCCTGCTAGTCTGTCAGGGC 540

Db 134421 CTTGGCCCCAGGGTTCAGCAGAGCTTCAGAGGTGGCCCCACCTGAGCCCCCACCAGGGGAGC 134480  
|||||  
QY 541 CTTGGCCCCAGGGTTCAGCAGAGCTTCAGAGGTGGCCCCACCTGAGCCCCCACCAGGGGAGC 600

Db 134481 AGTGTCTCTGTGCTTTTCTGTCATGCTTAGAGCATGTTCTTGGACATGGAATTT 134532  
|||||  
QY 601 AGTGTCTCTGTGCTTTTCTGTCATGCTTAGAGCATGTTCTTGGACATGGAATTT 652

RESULT 7

LOCUS BOVPBRIBP 821 bp mRNA MAM 05-AUG-1991  
DEFINITION Calf peripheral-type benzodiazepine receptor isoquinoline binding  
protein (PBR/IBP) mRNA, complete cds.

ACCESSION M64520

NID g163488

KEYWORDS benzodiazepine receptor isoquinoline binding protein.

SOURCE Calf adrenal gland, cDNA to mRNA.

ORGANISM Bos taurus

Eukaryotae; mitochondria eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovoidea;

Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 821)

AUTHORS

Parola,A.L., Stump,D.G., Pepper,D.J., Krueger,K.E., Regan,J.W. and  
Laird,H.E.II.

Cloning and expression of a pharmacologically unique bovine

peripheral-type benzodiazepine receptor isoquinoline binding

protein

JOURNAL J. Biol. Chem. 266, 14082-14087 (1991)

MEDLINE 91310699

FEATURES

source

Location/Qualifiers

1. 821

/organism="Bos taurus"

/db\_xref="taxon:9913"

/dev\_stage="calf"

/tissue\_type="adrenal gland"

/tissue\_lib="Of David Russell"

67. 806

/gene="PBR/IBP"

67. 576

/gene="PBR/IBP"

/codon\_start=1

/product="benzodiazepine receptor isoquinoline binding

protein"

/db\_xref="PID:g163489"

/translation="MAPPPVPAVGFLLPSLGGFLGAQYTRGEGFRWYASLQKPPWHP  
PRWILAPIWGTLYSAMGYSGYMIWKELGFSKEAVVPLGLYAGOLALNWAWPPLFFGT  
RQMGWALVDLLLTGGMAAATAMAHQVSPPAACLLYPYLAWLAFAGMLNRYMWDNQV  
RRSGRRRLSE"

801. 806

/gene="PBR/IBP"

polyA\_signal 124 a 278 c 255 g 164 t

BASE COUNT

ORIGIN

Query Match 44.2%; Score 288; DB 16; Length 821;  
Best Local Similarity 82.1%; Pred. No. 1.44e-196;  
Matches 381; Conservative 0; Mismatches 81; Indels 2; Gaps 2;

Db 144 CCGCGGAGAGGGTTCCGCTGGTACGGCAGCCTGCAGAAGCCCCCTGGCACCCGCCCG 203

|||||

QY 1 CCACGGCGAAGGTCTCCGCTGGTACGGCGGCCTGCAGAAGCCCTCGTGGCACCCGCCCA 60

Db 204 CTGGATTCTGGCTCCCATCTGGGGCACACTCTACTCGGCCATGGGATGTTCTCTACAT 263

|||||

QY 61 CTGGGTGCTGGGGCCTGTCTGGGCGACGCTCTACTCAGCCATGGSGTACGGCTCCTACCT 120

Db 264 GATCTGGAAGAGCTGGGGGGCTTCTCGAAGGAGGGGTGGTTCCTCGGCTCTACGC 323

|||||

QY 121 GGTCTGGAAGAGCTGGGAGGCTTCACAGAGAAGGCTGTGGTTCCTCGGCTCTACAC 180

Db 324 TGGGCAGCTGGCTCTGAACCTGGGCATGGCCTCCCTCTCTTCTTCGGGACTCGACAAATGGG 383

|||||

QY 181 TGGGCAGCTGGCCCTGAACCTGGGCATGGCCCCCATCTTCTTTGGTGGCCCTCTACAC 240

Db 384 CTGGGCCCTTGGTGGATCTCCTGTGACTGGCGGCATGGCAGCAGCCAGCCCATGGCCTG 443

|||||

QY 241 CTGGGCCCTTGGTGGATCTCCTGTGCTAGTGGGGCGGCGGCAGCCACTACCGTGGCCTG 300





QY 245 GCCTTGGTGATCTCCTGCTGGTGAGTGGGGCGGCGAGCCACTACCGTGGCTGGTAC 304

Db 416 CGAGTGAGCCACCGGCTGCGCGCTTGCTGTATCCTTACCTTACCTGGCCCTGGCTTGGCC 475

QY 305 CAGGTGAGCCCGCTGGCGCGCGCTGCTCTACCCCTACCTGGCTGGCTGGCTTCACG 364

Db 476 ACCATGCTCAACTACTATGATGGCGTGATAACTCTGGTCGGCGAGGGGGCTCCCGGCTC 535

QY 365 ACCACACTCAACTACTGCGTATGGCGGGACAAACCATGGCTGGCTGGGGGACGGCGGCTG 424

Db 536 ACAGAGTGAG 545

QY 425 CCAGAGTGAG 434

RESULT 10 MUSPTBR 856 bp mRNA ROD 27-MAR-1996

LOCUS Mouse mRNA for peripheral-type benzodiazepine receptor, complete cds.

DEFINITION D21207

ACCESSION 9484053

NID peripheral-type benzodiazepine receptor.

KEYWORDS Mus musculus cell-line erythro leukemia (library: lambda gt11) cDNA

SOURCE to mRNA.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 856)

AUTHORS Taketani,S., Kohno,H., Okuda,M., Furukawa,T. and Tokunaga,R.

TITLE Induction of peripheral-type benzodiazepine receptors during differentiation of mouse erythro leukemia cells. A possible involvement of these receptors in heme biosynthesis

JOURNAL J. Biol. Chem. 269, 2527-7531 (1994)

REFERENCE 2 (bases 1 to 856)

AUTHORS Taketani,S.

TITLE Direct Submission

JOURNAL Submitted (15-OCT-1993) to the DDBJ/EMBL/GenBank databases. Shigeru Taketani, Kansai Medical University, Dept. of Hygiene; 10-15 Fumizono-cho, Moriguchi, Osaka 570, Japan (Tel:06-992-1001(ex.2504), Fax:06-992-3522)

COMMENT Submitted (15-Oct-1993) to DDBJ by: Shigeru Taketani

Depertment of Hygiene

Kansai Medical University

1-Fumizonocho, Moriguchi

Osaka 570

Japan

Phone: 06-992-1001 x2504

Fax: 06-992-3522.

Location/Qualifiers

1. .856

/organism="Mus musculus"

/strain="BALB/c"

/sub\_species="domesticus"

/db\_xref="taxon:10090"

/cell\_line="erythro leukemia"

/clone\_lib="lambda gt11"

66. .575

/codon\_start=1

/product="peripheral-type benzodiazepine receptor"

/db\_xref="PID:d1005281"

/db\_xref="PID:g484054"

/translation="MPESWVPAVGLTVPSLGGFMGAYFVRGEGLRWYASLQKPSWHP PRWTLAPIWATLYSAMGYGYIVKELGGFTEDAMVPLGLYTGOLALNWAWPPIFFGA RQMGWALADLLLVSGVATATTLANHRSVPPAARLLYPYLANLAFATVLNYYVWRDMSG RRGGSRLAE"

BASE COUNT 150 a 254 c 253 g 199 t

ORIGIN

Query Match 42.0%; Score 274; DB 23; Length 856;

Best Local Similarity 79.5%; Pred. No. 1.71e-185;

Matches 369; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Db 147 GCGAGGGCCTCCGGTGGTATGCTAGCTGCAAGAAACCCCTTTGGCATCCGCTCGCTGG 206

QY 5 GCGAAGGTCTCCGCTGGTACGCGGCTGCAAGAGCCCTCGTGGACCGCCCACTGG 64

Db 207 AACTGGCTCCATCTGGCAACACTGTATTCAGCCATGGGGTATGGCTCTACATAGTC 266

QY 65 GTGCTGGGCCCTGTCTGGGCACGCTCTACTCAGCCATGGGGTACGCTCTACCTGGTC 124

Db 267 TGGAAAGAGCTGGGAGGTTTCACAGAGGAGCGCTATGTTCCCTTGGTCTCTACACTGGT 326

QY 125 TGGAAAGAGCTGGGAGGCTTCACAGAGAGGCTGTGTTCCCTTGGCTCTACACTGGG 184

Db 327 CAGCTGGCTCTGAACCTGGGCGTGGCCCCCATCTTCTTTGGTGGCGGCGAGATGGGCTGG 386

QY 185 CAGCTGGCCCTGAACCTGGGCATGGCCCCCATCTTCTTTGGTGGCGGCGAGATGGGCTGG 244

Db 387 GCCTTGGCCGATCTTCTGTGTAGTGGGGTGGCGACTGCCACAAACCTGGCTTGGCAC 446

QY 245 GCCTTGGTGGATCTCCTGTGTAGTGGGGCGGCGAGCCACTACCGTGGCTGGTAC 304

Db 447 CGAGTGAGCCCGCCGCTGCCGCTTGTGTATGCGTGTACCTTACCTGGCTGGCTTTTGGC 506

QY 305 CAGGTGAGCCCGCTGGCGCGCGCTGCTCTACCCCTACCTACCTGGCTGGCTTTCAGC 364

Db 507 ACCGTGCTCAACTACTATGTATGCGGTGATAACTCTGGCGGCGAGGGGCTCCCGGCTC 566

QY 365 ACCACACTCAACTACTCGTATGGCGGACACCATGGCTGGCGTGGGGACGGCGGCTG 424

Db 567 GCAGAGTGAAGGACCCAGCCATCAGGAATGCAGCCCTGCCAGC 610

QY 425 CCAGAGTGAGTGGCGCGGCCACCCAGGAGTGCAGCTGCACCAAGC 468

RESULT 11

LOCUS HSPBR3 645 bp DNA PRI 20-MAY-1994

DEFINITION Human peripheral benzodiazepine receptor gene, exon 3.

ACCESSION L21953

NID 9483404

KEYWORDS benzodiazepine receptor; peripheral benzodiazepine receptor.

SEGMENT 3 of 4

SOURCE Homo sapiens DNA.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 645)

AUTHORS Lin,D., Chang,Y.J., Strauss,J.F. and Miller,W.L.

TITLE The human peripheral benzodiazepine receptor gene: cloning and characterization of alternative splicing in normal tissues and in a patient with congenital lipoid adrenal hyperplasia

JOURNAL Genomics 18, 643-650 (1993)

MEDLINE 94140364

FEATURES Location/Qualifiers

source

1. .645

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/tissue\_type="placenta"

/tissue\_lib="placenta"

224. .362

exon

/product="peripheral benzodiazepine receptor"

BASE COUNT 138 a 177 c 192 g 138 t

ORIGIN

Query Match 19.6%; Score 128; DB 21; Length 645;

Best Local Similarity 98.6%; Pred. No. 6.34e-72;

Matches 140; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Db 223 GGTACGGCTCCTACCTGGTCTGGAAAGAGCTGGGAGGCTTCACAGAGAGGCTGTGGTT 282

QY 105 GGTACGGCTCCTACCTGGTCTGGAAAGAGCTGGGAGGCTTCACAGAGAGGCTGT-GGTT 163

Db 283 CCCCTGGGCTCTACACTGGGCAGCTGGCCCTGAACCTGGGCATGGCCCCCATCTTCTT- 341



QY 164 CCCCTGGGCTCTACACTGGGCAGCTGGCCCTGAACTGGGCATGGCCCCCATCTTCTTT 223

Db 342 GGTGCCCCGACAAATGGGCTGGG 363  
|||||

QY 224 GGTGCCCCGACAAATGGGCTGGG 245

RESULT 12  
LOCUS RATPBZR02 3434 bp DNA ROD 25-JAN-1993  
DEFINITION Rat peripheral-type benzodiazepine receptor gene exons 2-4,  
complete cds.  
ACCESSION M84221  
NID g206478  
KEYWORDS peripheral-type benzodiazepine receptor.  
SEGMENT 2 of 2  
SOURCE Rattus norvegicus (strain Wistar) Adult DNA.  
ORGANISM Rattus norvegicus  
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;  
Murinae; Rattus.  
REFERENCE 1 (bases 1 to 3434)  
AUTHORS Sprengel,R., Werner,P., Seeburg,P.H., Mukhin,A.G., Santi,R.M.,  
Grayson,D.R., Guidotti,A. and Krueger,K.E.  
TITLE Molecular Cloning and Expression of cDNA Encoding a Peripheral-type  
Benzodiazepine Receptor  
J. Biol. Chem. 264, 20415-20421 (1989)  
JOURNAL 90062173  
MEDLINE  
REFERENCE 2 (bases 1 to 3434)  
AUTHORS Casolotti,S.O., Pelala,G., Yakovlev,A.G., Csikos,T., Grayson,D.R.  
and Krueger,K.E.  
TITLE Structure of the rat gene encoding the mitochondrial benzodiazepine  
receptor  
Gene 121, 377-382 (1992)  
JOURNAL 93077057  
MEDLINE  
COMMENT Sequence M84221 starts with a Bam HI site, leaving a gap of about  
7kb between sequences.  
FEATURES  
source Location/Qualifiers  
1..3434  
/organism="Rattus norvegicus"  
/strain="Wistar"  
/db\_xref="taxon:10116"  
/dev\_stage="adult"  
order(M95864:1273..1462,1..1094)  
/number=1  
intron prim\_transcript 1..3350  
exon 1095..1305  
/note="putative"  
/number=2  
CDS join(1124..1305,1927..2065,2925..3113)  
/codon\_start=1  
/product="peripheral-type benzodiazepine receptor"  
/db\_xref="PID:g206480"  
/translation="MSQSWVPVAVGLTLVPSLGGFMGAYFVRGGLRWYASLQKPSWHP  
PRWTLAPIWGTLYSAMGYGSYIIWKELGGFTEAMVPLGLTGQLALNWAWPPIFFGA  
RQMGWALVDMLVSGVATATTLAHRVSPPAARLLYPYLAWLAFATMLNYYVWRDMSG  
RRGGSRLTE"  
intron 1306..1926  
exon 1927..2065  
intron 2066..2924  
exon 2925..3350  
BASE COUNT 713 a 872 c 991 g 858 t  
ORIGIN  
Query Match 17.6%; Score 115; DB 23; Length 3434;  
Best Local Similarity 90.8%; Pred. No. 3.97e-62;  
Matches 128; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
Db 1926 GGTATGGCTCCTACATAATCTGGAAGAGCTGGGAGGTTTCACAGAGGAGGCTATGGTTC 1985  
|||||

QY 105 GGTACGGCTCCTACCTGGTCTGGAAGAGCTGGGAGGCTTCACAGAGAAGGCTGTGGTTC 164

Db 1986 CCTTGGGTCTCTACACTGGTCACTGGCTCTGAACTGGGCATGGCCCCCATCTTCTTTG 2045  
|||||

QY 165 CCCTGGGCTCTACACTGGGCAGCTGGCCCTGAACTGGGCATGGCCCCCATCTTCTTTG 224

Db 2046 GTGCCCCGACAGATGGGCTGGG 2066  
|||||

QY 225 GTGCCCCGACAAATGGGCTGGG 245

RESULT 13  
LOCUS HSPBR2 487 bp DNA PRI 20-MAY-1994  
DEFINITION Human peripheral benzodiazepine receptor gene, exon 2.  
ACCESSION L21952  
NID g483403  
KEYWORDS benzodiazepine receptor; peripheral benzodiazepine receptor.  
SEGMENT 2 of 4  
SOURCE Homo sapiens DNA.  
ORGANISM Homo sapiens  
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 487)  
AUTHORS Lin,D., Chang,Y.J., Strauss,J.F. and Miller,W.L.  
TITLE The human peripheral benzodiazepine receptor gene: cloning and  
characterization of alternative splicing in normal tissues and in a  
patient with congenital lipoid adrenal hyperplasia  
Genomics 18, 643-650 (1993)  
JOURNAL 94140364  
MEDLINE  
FEATURES  
source Location/Qualifiers  
1..487  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/tissue\_type="placenta"  
/tissue\_lib="placenta"  
166..376  
/product="peripheral benzodiazepine receptor"  
BASE COUNT 67 a 168 c 156 g 96 t  
ORIGIN  
Query Match 16.3%; Score 106; DB 21; Length 487;  
Best Local Similarity 99.1%; Pred. No. 2.10e-55;  
Matches 107; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 272 CCACGGCGAGGGTCTCCGCTGGTACGCGCGCTGCAGAACCCCTCGTGGCACCCGCCCA 331  
|||||

QY 1 CCACGGCGAAGGTCTCCGCTGGTACGCGCGCTGCAGAACCCCTCGTGGCACCCGCCCA 60

Db 332 CTGGGTCTGGGCCCCTGTCTGGGGCAGCTCTACTACGCCATGGGTA 379  
|||||

QY 61 CTGGGTCTGGGCCCCTGTCTGGGGCAGCTCTACTACGCCATGGGTA 108

RESULT 14  
LOCUS I66494 7218 bp DNA PAT 23-DEC-1997  
DEFINITION Sequence 14 from patent US 5670367.  
ACCESSION I66494  
NID g2724471  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 7218)  
AUTHORS Dorner,F., Scheiflinger,F. and Falkner,F.Gunter.  
TITLE Recombinant fowlpox virus  
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;  
FEATURES  
source Location/Qualifiers  
1..7218  
/organism="unknown"  
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others  
ORIGIN  
Query Match 6.0%; Score 39; DB 18; Length 7218;

```
Best Local Similarity 0.3%; Pred. No. 1.82e-08;
Matches 1; Conservative 203; Mismatches 165; Indels 0; Gaps 0;

Db 1067 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1126
      :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
QY 13 TCTCCGCTGTACGGCGCCTGCAGAGCCCTCGTGGCACCGCCGCCCTGGGTGCTGGG 72

Db 1127 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1186
      :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
QY 73 CCCTGTCTGGGCACGCTCTACTCAGCCATGGGGTACGGCTCCTACCTGGTCTGGAAGA 132

Db 1187 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1246
      :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
QY 133 GCTGGGAGGCTTCACAGAGAAGGCTGTGTTCCCTGGCCTCTACACTGGGCAGTGGC 192

Db 1247 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1306
      :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
QY 193 CCTGAAC TGGGATGGCCCCCCTATCTTCTTTGGTGCCCGACAAATGGGCTTGGT 252

Db 1307 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1366
      :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
QY 253 GGATCTCCTGTGTCAGTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 312

Db 1367 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1426
      :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
QY 313 CCCGCTGGCGCCCGCCTGCTCTACCCCTACCTGGCCTGGCTGGCCTTACGACCACT 372

Db 1427 YYYYYYYG 1435
      : ::::
QY 373 CAACTACTG 381

RESULT 15
LOCUS CGU12420 1084 bp DNA ROD 14-DEC-1995
DEFINITION Cricetulus griseus mitochondrial benzodiazepine receptor (MBR)
          gene, partial cds.
ACCESSION U12420
NID 9529943
KEYWORDS Chinese hamster.
SOURCE Cricetulus griseus
        Cricetulus griseus
        Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
        Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
        Cricetinae; Cricetulus.
REFERENCE 1 (bases 41 to 1029)
AUTHORS Yakovlev,A.G., Ruffo,M., Jurka,J. and Krueger,K.E.
TITLE Comparison of repetitive elements in the third intron of human and
        rodent mitochondrial benzodiazepine receptor-encoding genes
JOURNAL Gene 155 (2), 201-205 (1995)
MEDLINE 95237610
REFERENCE 2 (bases 1 to 1084)
AUTHORS Krueger,K.E.
TITLE Direct Submission
JOURNAL Submitted (19-JUL-1994) Karl E. Krueger, Dept. of Cell Biology, USA
        Georgetown University School of Medicine, Washington, DC 20007, USA
FEATURES
    source
        1..1084
            /organism="Cricetulus griseus"
            /db_xref="taxon:10029"
            /clone_lib="PCR of genomic DNA"
            /cell_line="V79"
            1..1084
            /gene="MBR"
            <1..40
            /gene="MBR"
            /number=3
            /evidence=experimental
            /product="mitochondrial benzodiazepine receptor"
            join(<1..40,1030..>1084)
            /gene="MBR"
            /codon_start=2
            /product="mitochondrial benzodiazepine receptor"
            /db_xref="PID:g1039378"
```

```
/translation="WPIFFGARQMGWALADLLLVSGVATATTLA"
41..1029
/gene="MBR"
/number=3
/evidence=experimental
162..300
/rpt_family="B1"
340..422
/rpt_family="Alu-like"
657..736
/rpt_family="B1"
1030..>1084
/gene="MBR"
/number=4
/evidence=experimental
/product="mitochondrial benzodiazepine receptor"

BASE COUNT 253 a 279 c 302 g 250 t
ORIGIN

Query Match 5.7%; Score 37; DB 23; Length 1084;
Best Local Similarity 95.1%; Pred. No. 3.04e-07;
Matches 39; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 ATGGCCCCCATCTTCTTTGGTGCCCGGCAGATGGGCTGGG 41
  |||||
QY 205 ATGGCCCCCATCTTCTTTGGTGCCCGACAATGGGCTGGG 245

Search completed: Tue Jul 21 15:11:29 1998
Job time : 1092 secs.
```

**THIS PAGE BLANK (USPTO)**

\*\*\*\*\*  
[W][A][T][E][R][M][A][N]  
[S][E][Q][U][E][N][C][E]  
[A][L][I][G][N][M][E][N][T]  
[T][M]  
\*\*\*\*\*  
Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm  
Run on: Tue Jul 21 15:37:35 1998; MasPar time 830.75 Seconds  
Tabular output not generated. 1309.640 Million cell updates/sec  
Title: >US-09-047-652A-2  
Description: (1-652) from US09047652A.seq  
Perfect Score: 652  
N.A. Sequence: 1 CCACGGCGAGGCTCCGCT.....GTTCTTGGACATGGAATTT 652  
Comp: GGTGCCGCTCCAGAGCGGA.....CAAGAACCTTGTAACCTAAA  
Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0  
Searched: 457396 seqs, 834342348 bases x 2  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: embl54  
1:em\_ba 2:em\_htg 3:em\_hum1 4:em\_hum2 5:em\_in 6:em\_om  
7:em\_or 8:em\_ov 9:em\_pat 10:em\_pl 11:em\_ro 12:em\_vi  
Database: genbank106  
13:gb\_ba 14:gb\_htg 15:gb\_in 16:gb\_om 17:gb\_ov 18:gb\_pat  
19:gb\_ph 20:gb\_pl 21:gb\_pr1 22:gb\_pr2 23:gb\_ro 24:gb\_st  
25:gb\_sy 26:gb\_un 27:gb\_vi  
Statistics: Mean 10.376; Variance 5.505; scale 1.885  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES									
Result No.	Score	Query Match %	Length	DB	ID	Description	Pred. No.		
C	1	645	98.9	821	21	HUMHPBS	Human peripheral benzo	0.00e+00	
	2	542	83.1	639	21	HUMBENZA	Human peripheral benzo	0.00e+00	
	3	403	61.8	684	21	HSPBR4	Human peripheral benzo	1.05e-287	
	4	403	61.8	160559	14	HS1191B2	Human DNA sequence ***	1.05e-287	
	5	402	61.7	4258	21	HSU12421	Human mitochondrial be	6.58e-287	
	6	384	58.9	152843	14	HS526114	Human DNA sequence ***	1.53e-272	
	7	290	44.5	821	16	BOVPRIBP	Calf peripheral-type b	7.45e-198	
	8	284	43.6	626	23	MUSMPBR	Mouse peripheral-type	4.13e-193	
	9	278	42.6	781	23	RATPKBSX	Rat peripheral-type be	2.27e-188	
	10	276	42.3	856	23	MUSPTBR	Mouse mRNA for periphe	8.60e-187	
	11	128	19.6	645	21	HSPBR3	Human peripheral benzo	8.17e-72	
	12	115	17.6	3434	23	RATPTBZR02	Rat peripheral-type be	4.95e-62	
	13	108	16.6	487	21	HSPBR2	Human peripheral benzo	8.35e-57	
	14	39	6.0	7218	18	I66494	Sequence 14 from paten	1.90e-08	
	15	37	5.7	1084	23	CGU12420	Cricetulus grlseus mit	3.15e-07	

SUMMARIES

C	16	5.5	215	18	I28278	Sequence 5 from patent	1.26e-06
	17	5.5	1525	23	MMU12419	Mus musculus mitochond	1.26e-06
	18	5.4	215	18	I28278	Sequence 5 from patent	4.97e-04
	19	4.9	3084	13	S71770	carotenoid gene cluste	2.79e-04
	20	4.9	9298	13	SPCRT	R.sphaeroides crt gene	2.79e-04
	21	3.8	565	18	E04076	gDNA encoding envelope	1.73e+00
C	22	3.8	1284	22	HSAF000972	Homo sapiens calcium-a	1.73e+00
C	23	3.8	1284	22	AF022150	Homo sapiens intermedi	1.73e+00
C	24	3.8	1328	13	AEOAMOA	A.hydrophila amonabact	1.73e+00
	25	3.8	1684	21	HUMTTP	Human tristetraproline	1.73e+00
	26	3.8	1746	21	HUMGOS24A	H.sapiens zinc finger	1.73e+00
C	27	3.8	1982	22	AF033021	Homo sapiens intermed	1.73e+00
	28	3.8	2238	22	AF022797	Homo sapiens intermed	1.73e+00
C	29	3.8	3889	22	HUMGOS24B	Homo sapiens zinc fing	1.73e+00
	30	3.7	196	13	MSU77950	Mycobacterium smegmati	5.49e+00
	31	3.7	1120	13	AF005905	Mycobacterium smegmati	5.49e+00
	32	3.7	1804	22	HOSAWSCR1	Homo sapiens WSCR4 gen	5.49e+00
C	33	3.7	2926	13	S80432	ntxB, ntrC [Azospirill	5.49e+00
	34	3.7	4943	22	AB006629	Homo sapiens mRNA for	5.49e+00
	35	3.7	5350	23	RNU27319	Rattus norvegicus type	5.49e+00
	36	3.7	9483	23	MUSOPS	Mouse opsin (MOPS) gen	5.49e+00
C	37	3.7	134914	22	AC004542	Homo sapiens PAC clone	5.49e+00
C	38	3.5	69	18	I41362	Sequence 142 from pate	1.69e+01
	39	3.5	1697	21	HUMALPPC1	Human placental alkali	1.69e+01
C	40	3.5	2487	21	HUMALPLGC	Human alkaline phospho	1.69e+01
C	41	3.5	2688	21	HUMALPP	Human placental alkali	1.69e+01
C	42	3.5	5115	25	CVU89938	Cloning vector pSEAP2-	1.69e+01
	43	3.5	10068	13	RCCRATK	Rhodobacter capsulatus	1.69e+01
	44	3.5	119307	22	AC004134	Homo sapiens chromosom	1.69e+01
C	45	3.5	276261	14	HSAC000406	*** SEQUENCING IN PROG	1.69e+01

ALIGNMENTS

RESULT	1	HUMHPBS	821 bp	mRNA	PRI	20-DEC-1993
LOCUS		Human peripheral benzodiazepine receptor (hpbs) mRNA, complete cds.				
DEFINITION		M36035				
ACCESSION		9184333				
NID		peripheral benzodiazepine receptor.				
KEYWORDS		Human cDNA to mRNA.				
SOURCE		Homo sapiens				
ORGANISM		Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarhini; Hominidae; Homo.				
REFERENCE		1 (bases 1 to 821)				
AUTHORS		Riond, J., Mattel, M.G., Kaghad, M., Dumont, X., Guillemot, J.C., Le Fur, G., Caput, D. and Ferrara, P.				
TITLE		Molecular cloning and chromosomal localization of a human peripheral-type benzodiazepine receptor				
JOURNAL		Eur. J. Biochem. 195 (2), 305-311 (1991)				
MEDLINE		91146565				
REFERENCE		2 (bases 1 to 821)				
AUTHORS		Riond, J.				
TITLE		Direct Submission				
JOURNAL		Submitted (27-JUN-1990) J. Riond, Sanofi Elf Bio-Recherches, BP137, 31328 Labège Cedex, France				

FEATURES	Location/Qualifiers
source	1..821
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/cell_line="hystiocytic lymphoma monocyte-like cell line U937"
	/clone="p-hpBS11"
mRNA	<1..811
	/note="peripheral benzodiazepine receptor mRNA"
CDS	62..571
	/note="peripheral benzodiazepine receptor"
	/codon_start=1
	/db_xref="PID:g306883"
	/translation="MAPPPVPAMGFTLAPSLGFCVGRFVHGELRWYAGLOKPSWHP PHWVLGPVNGTLYSAMGYSLVWKELGFTKAVVPLGLYTGQLALNWAWPPIFFGA RQMGWALVDLLLVSGAAATTVAWYQVSLAARLLYPYLAWLAFATLTLCYVWRDNGH WHGRRRLPE"



polyA\_signal 800..805  
BASE COUNT 118 a 271 c 260 g 171 t 1 others  
ORIGIN Chromosome 22, map position q13.3.

Query Match 98.9%; Score. 645; DB 21; Length 821;  
Best Local Similarity 99.4%; Pred. No. 0.00e+00;  
Matches 648; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 139 CCACGGCGAGGGTCTCCGCTGGTACGCCGGCTGCAGAACCCCTCGTGGCACCCGCCCA 198  
QY 1 CCACGGCGAGGGTCTCCGCTGGTACGCCGGCTGCAGAACCCCTCGTGGCACCCGCCCA 60

Db 199 CTGGGTGCTGGGCCCTGTCTGGGGCACGCTCTACTACGCCATGGGTACGGCTCCTACCT 258  
QY 61 CTGGGTGCTGGGCCCTGTCTGGGGCACGCTCTACTACGCCATGGGTACGGCTCCTACCT 120

Db 259 GGTCTGGAAGAGCTGGGAGGCTTCACAGAGAAGGCTGTGTTCCCTGGGCTCTACAC 318  
QY 121 GGTCTGGAAGAGCTGGGAGGCTTCACAGAGAAGGCTGTGTTCCCTGGGCTCTACAC 180

Db 319 TGGGAGCTGGCCCTGAACCTGGGCATGGCCCCCATCTTCTTTGGTGGCCGACAAATGGG 378  
QY 181 TGGGAGCTGGCCCTGAACCTGGGCATGGCCCCCATCTTCTTTGGTGGCCGACAAATGGG 240

Db 379 CTGGGCTTGGTGGATCTCCTGTGTGTCAGTGGGGCGGCGNGCCACTACCGTGGCCTG 438  
QY 241 CTGGGCTTGGTGGATCTCCTGTGTGTCAGTGGGGCGGCGGAGCCACTACCGTGGCCTG 300

Db 439 GTACAGGTGAGCCCGCTGGCCGCCCGCTGCTCTACCCCTACCTGGCCTGGCCTT 498  
QY 301 GTACAGGTGAGCCCGCTGGCCGCCCGCTGCTCTACCCCTACCTGGCCTGGCCTT 360

Db 499 CGGACACACTCAACTACTGCGTATGGCGGGACAAACCATGGCTGGCATGGGGACGGCG 558  
QY 361 CAGGACACACTCAACTACTGCGTATGGCGGGACAAACCATGGCTGGCATGGGGACGGCG 420

Db 559 GCTGCCAGAGTGAGTGCCTGGCGGCGGCGGAGTGCAGTGCACAGAGTGCCATCAC 618  
QY 421 GCTGCCAGAGTGAGTGCCTGGCGGCGGCGGAGTGCAGTGCACAGAGTGCCATCAC 480

Db 619 GCTGTGATGTGGTGGCCGTACGCTTTCATGACCACTGGGCGCTGCTAGTCTGTAGGGC 678  
QY 481 GCTGTGATGTGGTGGCCGTACGCTTTCATGACCACTGGGCGCTGCTAGTCTGTAGGGC 540

Db 679 CTTGGCCCGAGGGTTCAGCAGAGCTTCAGAGGTTGCCCCACCTGAGCCCCACCCGGGAGC 738  
QY 541 CTTGGCCCGAGGGTTCAGCAGAGCTTCAGAGGTTGCCCCACCTGAGCCCCACCCGGGAGC 600

Db 739 AGTGCTCTGCTTTCTGTCATGCTTAGAGCATGTTCTTTGGAACATGGAATTT 790  
QY 601 AGTGCTCTGCTTTCTGTCATGCTTAGAGCATGTTCTTTGGAACATGGAATTT 652

RESULT 2  
LOCUS HUMBENZA 639 bp mRNA PRI 20-MAY-1994  
DEFINITION Human peripheral benzodiazepine receptor related mRNA sequence.  
ACCESSION L21950  
NID 9483401  
KEYWORDS benzodiazepine receptor; peripheral benzodiazepine receptor.  
SOURCE Homo sapiens cdna to mRNA.  
ORGANISM Homo sapiens  
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 639)  
AUTHORS Lin,D., Chang,Y.J., Strauss,J.F. and Miller,W.L.  
TITLE The human peripheral benzodiazepine receptor gene: cloning and characterization of alternative splicing in normal tissues and in a patient with congenital lipoid adrenal hyperplasia  
JOURNAL Genomics 18 (3), 643-650 (1993)  
MEDLINE 94140364  
FEATURES Location/Qualifiers  
1..639  
/organism="Homo sapiens"

/db\_xref="taxon:9606"  
/tissue\_type="placenta"  
161..469  
/codon\_start=1  
/product="peripheral benzodiazepine receptor related protein"  
/db\_xref="PID:g488420"  
/translation="MAPHLLWCPTNGLGGSPAGQWGGSHYRGLVPGEPAGRPPAL  
PLPGLAGLRDHTQLLRMAGQPLAWGTAAARVSARPTRDCSCTSRCHACDVAVTLS  
"  
polyA\_site 639  
BASE COUNT 100 a 201 c 200 g 138 t  
ORIGIN

Query Match 83.1%; Score 542; DB 21; Length 639;  
Best Local Similarity 99.5%; Pred. No. 0.00e+00;  
Matches 545; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 61 GGTACGGCTCCTACCTGGTCTGGAAGAGCTGGGAGGCTTCACAGAGAAGGCTGTGTTT 120  
QY 105 GGTACGGCTCCTACCTGGTCTGGAAGAGCTGGGAGGCTTCACAGAGAAGGCTGTGTTT 164

Db 121 CCCTGGGCTCTACACTGGGAGCTGGCCCTGAACCTGGGATGGCCCCCATCTTCTTTT 180  
QY 165 CCCTGGGCTCTACACTGGGAGCTGGCCCTGAACCTGGGATGGCCCCCATCTTCTTTT 224

Db 181 GTGCCCCACAAATGGGCTGGGCTTGGTGGATCTCCTGCTGTCAGTGGGGCGGCGCAG 240  
QY 225 GTGCCCCACAAATGGGCTGGGCTTGGTGGATCTCCTGCTGTCAGTGGGGCGGCGCAG 284

Db 241 CCACTACCGTGGCTTGGTACCGAGTGGCCCGCTGGCCCGCTGCTTACCCCTACC 300  
QY 285 CCACTACCGTGGCTTGGTACCGAGTGGCCCGCTGGCCCGCTGCTTACCCCTACC 344

Db 301 TGGCCTGGCTGGCCTTCGCGACCACTCAACTACTGCGTATGGCGGACAAACCATGGCT 360  
QY 345 TGGCCTGGCTGGCCTTCACGACCACTCAACTACTGCGTATGGCGGACAAACCATGGCT 404

Db 361 GGCATGGGGACGGCGGCTGCCAGAGTGAAGTCCCGGCCACCAGGACTGCAGTGCAC 420  
QY 405 GGCATGGGGACGGCGGCTGCCAGAGTGAAGTCCCGGCCACCAGGACTGCAGTGCAC 464

Db 421 CAGCAGGTGCCATCAGCCTTGTGATGTGGTGGCGCTCAGCCTTTCATGACCACTGGGCT 480  
QY 465 CAGCAGGTGCCATCAGCCTTGTGATGTGGTGGCGCTCAGCCTTTCATGACCACTGGGCT 524

Db 481 GCTAGTCTGTAGGGCCTTGGCCCGAGGTCAGAGAGCTTCAGAGGTTGCCCCACCTGA 540  
QY 525 GCTAGTCTGTAGGGCCTTGGCCCGAGGTCAGAGAGCTTCAGAGGTTGCCCCACCTGA 584

Db 541 GCCCCACCCGGGAGCAGTGTCTGCTTCTGTCATGCTTAGAGCATGTTCTTGAACA 600  
QY 585 GCCCCACCCGGGAGCAGTGTCTGCTTCTGTCATGCTTAGAGCATGTTCTTGAACA 644

Db 601 TGAATTT 608  
QY 645 TGAATTT 652

RESULT 3  
LOCUS HSPBR4 684 bp DNA PRI 20-MAY-1994  
DEFINITION Human peripheral benzodiazepine receptor gene, exon 4.  
ACCESSION L21954  
NID 9483405  
KEYWORDS benzodiazepine receptor; peripheral benzodiazepine receptor.  
SEGMENT 4 of 4  
SOURCE Homo sapiens DNA.  
ORGANISM Homo sapiens  
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 684)  
AUTHORS Lin,D., Chang,Y.J., Strauss,J.F. and Miller,W.L.  
TITLE The human peripheral benzodiazepine receptor gene: cloning and



characterization of alternative splicing in normal tissues and in a patient with congenital lipoid adrenal hyperplasia  
Genomics 18, 643-650 (1993)  
94140364

FEATURES  
source

Location/Qualifiers

1. .684  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/tissue\_type="placenta"  
/tissue\_lib="placenta"  
join(L21952:195. .376,L21953:224. .362,164. .352)  
/codon\_start=1  
/product="peripheral benzodiazepine receptor"  
/db\_xref="PID:g488425"  
PHWLGVPWGTLYSAMGYGSYLWKELGGFTEKAVGSPGLHWAAGPELGMAPHLIGA  
ROMGWALVDLLLVSGAAAATTVAWYQVSPLAARLLPYLAWLAFATTLNYCVWRDNHG  
WHGGRRLPE"  
164. .602  
/product="peripheral benzodiazepine receptor"

BASE COUNT 113 a 200 c 225 g 146 t  
ORIGIN

Query Match 61.8%; Score 403; DB 21; Length 684;  
Best Local Similarity 99.3%; Pred. No. 1.05e-287;  
Matches 406; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 163 GGCCTTGGTGATCTCCTGCTGGTCAGTGGGGCGGCGGCAGCCACTACCGTGGCCTGGTA 222  
|||||  
QY 244 GGCCTTGGTGATCTCCTGCTGGTCAGTGGGGCGGCGGCAGCCACTACCGTGGCCTGGTA 303  
|||||  
Db 223 CCAGGTGAGCCCGCTGGCCGCCCTGCTCTACCCCTACCTGGCCTGGCTGGCCTTCGC 282  
|||||  
QY 304 CCAGGTGAGCCCGCTGGCCGCCCTGCTCTACCCCTACCTGGCCTGGCTGGCCTTCAC 363  
|||||  
Db 283 GACCACACTCAACTACTCGTATGGCGGGACACCATGGCTGGCATGGGGACGGCGGCT 342  
|||||  
QY 364 GACCACACTCAACTACTCGTATGGCGGGACACCATGGCTGGCTGGGGACGGCGGCT 423  
|||||  
Db 343 GCCAGAGTGAGTGCCCGGCCACCAGGGACTGCAGCTGCACCAGCAGGTGCCATCAGCT 402  
|||||  
QY 424 GCCAGAGTGAGTGCCCGGCCACCAGGGACTGCAGCTGCACCAGCAGGTGCCATCAGCT 483  
|||||  
Db 403 TGTGATGTGGTGGCCCTCACGCTTTTCATGACCCTGGGCGCTGCTAGTCTGTGAGGGCCTT 462  
|||||  
QY 484 TGTGATGTGGTGGCCCTCACGCTTTCATGACCCTGGGCGCTGCTAGTCTGTGAGGGCCTT 543  
|||||  
Db 463 GGGCCAGGGGTCAGCAGAGCTTCAGAGGTTGCCCCACCTGAGCCCCCAGCCGGGAGCAGT 522  
|||||  
QY 544 GGGCCAGGGGTCAGCAGAGCTTCAGAGGTTGCCCCACCTGAGCCCCCAGCCGGGAGCAGT 603  
|||||  
Db 523 GTCCTGTGCTTTCTGATGCTTAGAGCATGTTCTTGGAAACATGGAATTT 571  
|||||  
QY 604 GTCCTGTGCTTTCTGATGCTTAGAGCATGTTCTTGGAAACATGGAATTT 652  
|||||

RESULT 4  
LOCUS HS1191B2 160559 bp DNA HTG 25-MAR-1998  
DEFINITION Human DNA sequence \*\*\* SEQUENCING IN PROGRESS \*\*\* from clone 1191B2; HTGS phase 1.  
ACCESSION AL022237  
NID 92995178  
KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Sulston,J.  
TITLE Direct Submission  
JOURNAL Submitted (23-MAR-1998) Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

COMMENT

IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continues. The sequence may be contaminated with foreign sequence from E.coli, yeast, vector, phage etc. Order of segments is not known; 800 n's separate segments. Unfinished sequence: bk1191B2 Contig\_ID: 01719 acc= Length: 2099 bp Unfinished sequence: bk1191B2 Contig\_ID: 01607 acc= Length: 5280 bp Unfinished sequence: bk1191B2 Contig\_ID: 01221 acc= Length: 1033 bp Unfinished sequence: bk1191B2 Contig\_ID: 01006 acc= Length: 1183 bp Unfinished sequence: bk1191B2 Contig\_ID: 01520 acc= Length: 1158 bp Unfinished sequence: bk1191B2 Contig\_ID: 01212 acc= Length: 1243 bp Unfinished sequence: bk1191B2 Contig\_ID: 00949 acc= Length: 6912 bp Unfinished sequence: bk1191B2 Contig\_ID: 00953 acc= Length: 3574 bp Unfinished sequence: bk1191B2 Contig\_ID: 00946 acc= Length: 4065 bp Unfinished sequence: bk1191B2 Contig\_ID: 00376 acc= Length: 2336 bp Unfinished sequence: bk1191B2 Contig\_ID: 00587 acc= Length: 2115 bp Unfinished sequence: bk1191B2 Contig\_ID: 00546 acc= Length: 3119 bp Unfinished sequence: bk1191B2 Contig\_ID: 01145 acc= Length: 3772 bp Unfinished sequence: bk1191B2 Contig\_ID: 01831 acc= Length: 5422 bp Unfinished sequence: bk1191B2 Contig\_ID: 01683 acc= Length: 5648 bp Unfinished sequence: bk1191B2 Contig\_ID: 01825 acc= Length: 9110 bp Unfinished sequence: bk1191B2 Contig\_ID: 01829 acc= Length: 12009 bp Unfinished sequence: bk1191B2 Contig\_ID: 00990 acc= Length: 14198 bp Unfinished sequence: bk1191B2 Contig\_ID: 00103 acc= Length: 17541 bp Unfinished sequence: bk1191B2 Contig\_ID: 01695 acc= Length: 20537 bp Unfinished sequence: bk1191B2 Contig\_ID: 00329 acc= Length: 5642 bp Unfinished sequence: bk1191B2 Contig\_ID: 00606 acc= Length: 8010 bp Unfinished sequence: bk1191B2 Contig\_ID: 01248 acc= Length: 1799 bp Unfinished sequence: bk1191B2 Contig\_ID: 00095 acc= Length: 4354 bp.

\*\*\* WARNING: Phase 1 High Throughput Genome Sequence \*\*\*

\*\*\* This sequence is unfinished. When sequencing is complete, \* the sequence data presented in this record will be replaced \*by a single finished sequence with the same accession number.

FEATURES  
source

1. 160559  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="22"  
/clone="1191B2"

BASE COUNT 33840 a 38651 c 38486 g 31167 t 18415 others  
ORIGIN

Query Match 61.8%; Score 403; DB 14; Length 160559;  
Best Local Similarity 99.3%; Pred. No. 1.05e-287;  
Matches 406; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 71240 AAATTCCATGTTCCAAGAACATGCTCTAAGCATGCACAAAGCACAGGACACTGCTCCCG 71299  
|||||  
Cp 652 AAATTCCATGTTCCAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCG 593  
|||||  
Db 71300 GTGGGGCTCAGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGSCCAAGCCCTGAC 71359  
|||||  
Cp 592 GTGGGGCTCAGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGSCCAAGCCCTGAC 533  
|||||  
Db 71360 AGACTAGCACGCCAGTGGTCATGAAGCGTGACGGCCACCACATCATCAAGCGTGATGGC 71419  
|||||  
Cp 532 AGACTAGCAGGCCAGTGGTCATGAAGCGTGACGGCCACCACATCATCAAGCGTGATGGC 473  
|||||  
Db 71420 ACCTGCTGGTCAGCTGCAGTCCCTGGTGGGCGGGCAGCTCACTCTGGCAGCCCGCTCC 71479  
|||||  
Cp 472 ACCTGCTGGTCAGCTGCAGTCCCTGGTGGGCGGGCAGCTCACTCTGGCAGCCCGCTCC 413  
|||||  
Db 71480 CCCACGCCAGCCATGGTTGTCCCGCCATACGCAGTAGTTGAGTGTGTCGCGAAGGCCAG 71539  
|||||  
Cp 412 CCCACGCCAGCCATGGTTGTCCCGCCATACGCAGTAGTTGAGTGTGTCGTAAGGCCAG 353  
|||||

Db 71540 CCAGGCCAGGTAGGGGTAGAGCAGCGGGCGGCAGCGGGCTCACCTGGTACCAGGCCAC 71599  
|||||  
Cp 352 CCAGGCCAGGTAGGGGTAGAGCAGCGGGCGGCAGCGGGCTCACCTGGTACCAGGCCAC 293  
Db 71600 GGTAGTGGCTGCCCGCCCCCTACTGACCAGCAGGAGATCCACCAAGGCC 71648  
|||||  
Cp 292 GGTAGTGGCTGCCCGCCCCCTACTGACCAGCAGGAGATCCACCAAGGCC 244

RESULT 5  
LOCUS HSU12421 4258 bp DNA PRI 14-DEC-1995  
DEFINITION Human mitochondrial benzodiazepine receptor (MBR) gene, complete cds.

ACCESSION U12421  
NID 9529945  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 2080 to 3683)  
AUTHORS Yakovlev,A.G., Ruffo,M., Jurka,J. and Krueger,K.E.  
TITLE Comparison of repetitive elements in the third intron of human and rodent mitochondrial benzodiazepine receptor-encoding genes  
JOURNAL Gene 155 (2), 201-205 (1995)  
MEDLINE 95237610  
REFERENCE 2 (bases 1 to 4258)  
AUTHORS Krueger,K.E.  
TITLE Direct Submission  
JOURNAL Submitted (19-JUL-1994) Karl E. Krueger, Dept. of Cell Biology, Georgetown University School of Medicine, Washington, DC 20007, USA  
FEATURES  
source 1..4258  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="22q13.3"  
intron <1..99  
/number=1  
/evidence=experimental  
exon 100..310  
/number=2  
/evidence=experimental  
/product="mitochondrial benzodiazepine receptor"  
5'UTR 100..128  
gene 129..3872  
/gene="MBR"  
CDS join(129..310,1941..2079,3684..3872)  
/gene="MBR"  
/codon\_start=1  
/evidence=experimental  
/product="mitochondrial benzodiazepine receptor"  
/db\_xref="PID:g529946"  
/translation="MAPPWVPAMGFTLAPSLGCFVGSRFVHGELRWYAGLQKPSWHP  
PHWLVGPVWGTLYSAMGYGSYLVWKELGFTTEKAVVPLGLYTGQLALNWAWPPIFFGA  
RQMGWALVDLLLVSGAAATTVAWYQVSPLAARLLYPYLAWLAFITTLNYCVWRDNG  
WRGRRRLPE"  
intron 311..1940  
/gene="MBR"  
/number=2  
/evidence=experimental  
exon 1941..2079  
/gene="MBR"  
/number=3  
/evidence=experimental  
/product="mitochondrial benzodiazepine receptor"  
intron 2080..3683  
/gene="MBR"  
/number=3  
/evidence=experimental  
repeat\_region 2401..3306  
/note="4 Alus"  
repeat\_region /rpt\_family="Alu"  
3497..3581

exon /rpt\_family="MIR"  
3684..4122  
/number=4  
/evidence=experimental  
/product="mitochondrial benzodiazepine receptor"  
3873..4122  
BASE COUNT 794 a 1246 c 1221 g 997 t  
ORIGIN

Query Match 61.7%; Score 402; DB 21; Length 4258;  
Best Local Similarity 99.8%; Pred. No. 6.58e-287;  
Matches 408; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 3683 GGCCTTGGTGGATCTCCTGCTGCTAGTGGGCGGCGGAGCCACTACCGTGGCCTGGTA 3742  
|||||  
QY 244 GGCCTTGGTGGATCTCCTGCTGCTAGTGGGCGGCGGAGCCACTACCGTGGCCTGGTA 303  
Db 3743 CCAGGTGAGCCCGCTGGCGCGCGCTGCTTACCCCTACCTGGCCTGGCTTTCAC 3802  
|||||  
QY 304 CCAGGTGAGCCCGCTGGCGCGCGCTGCTTACCCCTACCTGGCCTGGCTTTCAC 363  
Db 3803 GACCACACTCAACTACTGCGTATGCGGGGACAAACCATGGCTGGGCGGCGGCT 3862  
|||||  
QY 364 GACCACACTCAACTACTGCGTATGCGGGGACAAACCATGGCTGGGCGGCGGCT 423  
Db 3863 GCCAGAGTGAGTCCCGGCCACACAGGAGTGCAGCTGCACCAGAGGTGCCATCAGCT 3922  
|||||  
QY 424 GCCAGAGTGAGTCCCGGCCACACAGGAGTGCAGCTGCACCAGAGGTGCCATCAGCT 483  
Db 3923 TGTGATGGTGGTGGCGGCTCAGCGCTTCATGACCACTGGGCGCTGCTAGTCTGTGAGGGCCCTT 3982  
|||||  
QY 484 TGTGATGGTGGTGGCGGCTCAGCGCTTCATGACCACTGGGCGCTGCTAGTCTGTGAGGGCCCTT 543  
Db 3983 GGCCAGGGGTACAGCAGAGCTTACAGAGTGGCCCCAC-TGAGCCCCACCCGGGAGCAGT 4041  
|||||  
QY 544 GGCCAGGGGTACAGCAGAGCTTACAGAGTGGCCCCACCTGAGCCCCACCCGGGAGCAGT 603  
Db 4042 GTCCTGTGCTTCTGTCATGCTTAGAGCATGTTCTTGAACATGGAATTT 4090  
|||||  
QY 604 GTCCTGTGCTTCTGTCATGCTTAGAGCATGTTCTTGAACATGGAATTT 652

RESULT 6  
LOCUS HS526I14 152843 bp DNA HTG 13-JAN-1997  
DEFINITION Human DNA sequence \*\*\* SEQUENCING IN PROGRESS \*\*\* from clone 526I14; HTGS phase 1.  
ACCESSION Z82214  
NID g1666406  
KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 152843)  
AUTHORS Buck,D.  
TITLE Direct Submission  
JOURNAL Submitted (06-NOV-1996) Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquires: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
COMMENT IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continues. The sequence may be contaminated with foreign sequence from E.coli, yeast, vector, phage etc. Order of segments is not known; 800 n's separate segments. Unfinished sequence: dj526I14 Contig\_ID: 00003 Length: 15560 bp Unfinished sequence: dj526I14 Contig\_ID: 00229 Length: 791 bp Unfinished sequence: dj526I14 Contig\_ID: 00240 Length: 810 bp Unfinished sequence: dj526I14 Contig\_ID: 00248 Length: 750 bp Unfinished sequence: dj526I14 Contig\_ID: 00250 Length: 763 bp Unfinished sequence: dj526I14 Contig\_ID: 00263 Length: 794 bp Unfinished

sequence: dJ526114 Contig\_ID: 00278 Length: 1221 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00281 Length: 1174 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00287 Length: 1142 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00295 Length: 1145 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00308 Length: 826 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00327 Length: 1117 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00342 Length: 1668 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00351 Length: 947 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00359 Length: 1163 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00369 Length: 1092 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00393 Length: 1262 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00408 Length: 2766 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00423 Length: 1714 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00435 Length: 2050 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00453 Length: 2436 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00547 Length: 3304 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00574 Length: 4133 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00610 Length: 4473 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00647 Length: 4540 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00680 Length: 3551 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00752 Length: 3422 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00793 Length: 4738 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00831 Length: 4862 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00870 Length: 2306 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00944 Length: 7105 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00995 Length: 6767 bp Unfinished  
sequence: dJ526114 Contig\_ID: 01058 Length: 4554 bp Unfinished  
sequence: dJ526114 Contig\_ID: 01131 Length: 8507 bp Unfinished  
sequence: dJ526114 Contig\_ID: 01204 Length: 9125 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00899 Length: 10714 bp Unfinished  
sequence: dJ526114 Contig\_ID: 00896 Length: 751 bp  
\*\*\*  
\*\*\* WARNING: Phase 1 High Throughput Genome Sequence \*\*\*  
\*\*\*  
\* This sequence is unfinished. When sequencing is complete,  
\* the sequence data presented in this record will be replaced  
\* by a single finished sequence with the same accession number.

FEATURES  
source  
1. .152843  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="526114"  
/chromosome="22"  
BASE COUNT 28005 a 34564 c 33570 g 27611 t 29093 others  
ORIGIN

Query Match 58.9%; Score 384; DB 14; Length 152843;  
Best Local Similarity 98.1%; Pred. No. 1.53e-272;  
Matches 404; Conservative 0; Mismatches 5; Indels 3; Gaps 3;  
Db 134121 GGCCTGGTGATCTCCGCTGGTNCANGTGGGCGCGNNGCCACTACCGTGGCCTGG 134180  
QY 244 GGCCTGGTGATCTCCTGCTGGT-CA-GTGGGGCGCGGCAGCCACTACCGTGGCCTGG 301  
Db 134181 TACCAGGTGAGCCCGCTGSCCGCCCGCTGCTATACCCCTACCTGGCCTGGCTGGCCTTC 134240  
QY 302 TACCAGGTGAGCCCGCTGSCCGCCCGCTGCTATACCCCTACCTGGCCTGGCTGGCCTTC 361  
Db 134241 ACGACCACACTCAACTACTGCGTATGGCGGGACAACCATGGCTGGCGTGGGGACGGCGG 134300  
QY 362 ACGACCACACTCAACTACTGCGTATGGCGGGACAACCATGGCTGGCGTGGGGACGGCGG 421  
Db 134301 CTGCCAGAGTGAGTGCCCGGCCCCACCAGGGACTGCAGTGCACACAGCAGGTGCCATCACG 134360  
QY 422 CTGCCAGAGTGAGTGCCCGGCCCCACCAGGGACTGCAGTGCACACAGCAGGTGCCATCACG 481  
Db 134361 CTTGNTGATGTGGTGGCCGCTCACGCTTTTCATGACCACTGGGCCTGCTAGTCTGCAGGGC 134420  
QY 482 CTTG-TGATGTGGTGGCCGCTCACGCTTTTCATGACCACTGGGCCTGCTAGTCTGCAGGGC 540  
Db 134421 CTTGGCCACAGGGGTGAGCAGAGCTTCAGAGGTGGCCCGCCACCTGAGCCCCCACCAGGGGAGC 134480  
QY 541 CTTGGCCACAGGGGTGAGCAGAGCTTCAGAGGTGGCCCGCCACCTGAGCCCCCACCAGGGGAGC 600

Db 134481 AGTGTCTCTGTGCTTTCTGTCATGCTTAGAGCATGTTCTTGGAACATGGAATTT 134532  
QY 601 AGTGTCTCTGTGCTTTCTGTCATGCTTAGAGCATGTTCTTGGAACATGGAATTT 652  
RESULT 7  
LOCUS BOVBRIBP 821 bp mRNA MAM 05-AUG-1991  
DEFINITION Calf peripheral-type benzodiazepine receptor isoquinoline binding protein (PBR/IBP) mRNA, complete cds.  
ACCESSION M64520  
NID g163488  
KEYWORDS benzodiazepine receptor isoquinoline binding protein.  
SOURCE Calf adrenal gland, cDNA to mRNA.  
ORGANISM Bos taurus  
Eukaryotae; mitochondria eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.  
REFERENCE 1 (bases 1 to 821)  
AUTHORS Parola,A.L., Stump,D.G., Pepper,D.J., Krueger,K.E., Regan,J.W. and Laird,H.E.II.  
TITLE Cloning and expression of a pharmacologically unique bovine peripheral-type benzodiazepine receptor isoquinoline binding protein  
JOURNAL J. Biol. Chem. 266, 14082-14087 (1991)  
MEDLINE 91310699  
FEATURES  
Location/Qualifiers  
1..821  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/dev\_stage="calf"  
/tissue\_type="adrenal gland"  
/tissue\_lib="of David Russell"  
67..806  
/gene="PBR/IBP"  
67..576  
/gene="PBR/IBP"  
/codon\_start=1  
/product="benzodiazepine receptor isoquinoline binding protein"  
/db\_xref="PID:g163489"  
/translation="MAPWPVPAVGFTLLPSLGGFLGAQYTRGEGFRWYASLQKPPWHP PRWILAPIWGTLYSAMGYCSYMIWKELGGFSKEAVVPLGLYAGQLALNWAWPPLFFGT RQMGWALVDLLLTGGMAAATAMAHQVSPPAACLLYPYLAWLAFAGMLNRYMWQDNQV RRSGRRLSE"  
801..806  
polyA\_signal  
/gene="PBR/IBP"  
BASE COUNT 124 a 278 c 255 g 164 t  
ORIGIN  
Query Match 44.5%; Score 290; DB 16; Length 821;  
Best Local Similarity 82.3%; Pred. No. 7.45e-198;  
Matches 382; Conservative 0; Mismatches 80; Indels 2; Gaps 2;  
Db 144 CCGCGGAGAGGGTTCCGCTGGTACGCCAGCCTGCAGAAGCCCCCGTGGCACCCGCCCG 203  
QY 1 CCACGGGAGGGTCTCCGCTGGTACGCCGGCTGCAGAAGCCCTCGTGGCACCCGCCCCA 60  
Db 204 CTGGATTCTGGCTCCCATCTGGGGCACACTCTACTCGGCCATGGGGTATGGTTCCTACAT 263  
QY 61 CTGGGTGCTGGGCCTGTCTGGGGCACGCTCTACTCAGCCATGGGGTACGGCTCCTACCT 120  
Db 264 GATCTGGAAGAGCTGGGGGCTTCTCGAAGAGGCGGCTGGTTCCTCCCTGGGCCTCTACGC 323  
QY 121 GGTCTGGAAGAGCTGGGAGGCTTCACAGAGAAGGCTGTGGTTCCTCCCTGGGCCTCTACAC 180  
Db 324 TGGCAGCTGGCTCTGAACCTGGGCATGGCCTCCCTCTCTTCTTCGGGCACCTCGACAAATGG 383  
QY 181 TGGCAGCTGGCCCTGAACCTGGGCATGGCCCCCCTCTCTTTTGGTCCCGACAAATGG 240  
Db 384 CTGGGCCTTGGTGGATCTCCTGCTGACTGGCGGCATGGCAGCAGCCACGGCCATGGCCTG 443  
QY 241 CTGGGCCTTGGTGGATCTCCTGCTGGTCAAGTGGGGCGGGCAGCCACTACCGTGGCCTG 300





QY 245 GCCTTGGTGATCTCCTGCTGGTCACTGGGGGGGGCAGCCACTACCGTGGCCTGGTAC 304

Db 416 CGAGTGAGCCACCGGCTGCCCGCTTGCTGTATCCTTACCTGGCCCTGGCTTTGCC 475

QY 305 CAGGTGAGCCGCTGGCGCCGCTCTACTACCCCTACCTACCTGGCCTGGCTTTCACG 364

Db 476 ACCATGCTCAACTACTATGTATGGCGTGATAACTCTGGTCGGCGAGGGGGCTCCCGGCTC 535

QY 365 ACCACACTCAACTACTGCGTATGGCGGGGACAACCATGGCTGGCTGGGGGACGGCGGCTG 424

Db 536 ACAGAGTGAG 545

QY 425 CCAGAGTGAG 434

RESULT 10 MUSPTBR 856 bp mRNA ROD 27-MAR-1996

LOCUS Mouse mRNA for peripheral-type benzodiazepine receptor, complete cds.

DEFINITION D21207

ACCESSION 9484053

NID peripheral-type benzodiazepine receptor.

KEYWORDS Mus musculus cell-line erythroleukemia (library: lambda gt11) cDNA

SOURCE to mRNA.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 856)

AUTHORS Taketani,S., Kohno,H., Okuda,M., Furukawa,T. and Tokunaga,R.

TITLE Induction of peripheral-type benzodiazepine receptors during differentiation of mouse erythroleukemia cells. A possible involvement of these receptors in heme biosynthesis

JOURNAL J. Biol. Chem. 269, 2527-7531 (1994)

REFERENCE 2 (bases 1 to 856)

AUTHORS Taketani,S.

TITLE Direct Submission

JOURNAL Submitted (15-OCT-1993) to the DDBJ/EMBL/GenBank databases. Shigeru Taketani, Kansai Medical University, Dept. of Hygiene; 10-15 Fumizono-cho, Moriguchi, Osaka 570, Japan (Tel:06-992-1001(ex.2504), Fax:06-992-3522)

COMMENT Submitted (15-Oct-1993) to DDBJ by: Shigeru Taketani

Department of Hygiene

Kansai Medical University

1-Fumizonocho, Moriguchi

Osaka 570

Japan

Phone: 06-992-1001 x2504

Fax: 06-992-3522.

Location/Qualifiers

1. .856

/organism="Mus musculus"

/strain="BALB/c"

/sub\_species="domesticus"

/db\_xref="taxon:10090"

/cell\_line="erythroleukemia"

/clone\_lib="lambda gt11"

66. .575

/codon\_start=1

/product="peripheral-type benzodiazepine receptor"

/db\_xref="PID:d1005281"

/db\_xref="PID:g484054"

/translation="MPESWVPVAVGLTLVPSLGGFMGAYFVRGEGLRWYASLQKPSWHP

PRWTLAPIWATLYSAMGYGIYWKELGGFTEDAMVPLGLYTGQALNWAWPPIFFGA

ROMGWALADLLLVSGVATATTLAWHRVSPPAARLLPYLAWLAFATVLNYYVWRDMSG

RRGSRLAE"

BASE COUNT 150 a 254 c 253 g 199 t

ORIGIN

Query Match 42.3%; Score 276; DB 23; Length 856;

Best Local Similarity 79.7%; Pred. No. 8.60e-187;

Matches 370; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Db 147 GCGAGGGCCCTCCGGTGGTATGTAGTCTTGCAAGAAACCCTCTTGGCATCCGCCCTCGCTGG 206

QY 5 GCGAGGGTCTCCGCTGGTACGCGGGCCTGCAGAAGCCCTCGTGGCACCGGCCCACTGG 64

Db 207 ACACGTGGCTCCCATCTGCGGCAACACTGTATTACGCCCATGSGGTATGGCTCTACATAGTC 266

QY 65 GTGCTGGCCCTGTCTGGGGCACGCTCTACTCAGCCCATGSGGTACGGCTCCTACCTGGTC 124

Db 267 TGGAAAGAGCTGGGAGGTTTCACAGAGGACGCTATGGTTCCTTGGGTCTCTACACTGGT 326

QY 125 TGGAAAGAGCTGGGAGGCTTCACAGAGAAGGCTGTGGTTCCTGGCCCTCTACACTGGG 184

Db 327 CAGCTGGCTCTGAACCTGGGCGTGGCCCCCCCACATCTTCTTTGGTGCCCCGCGAGATGGGCTG 386

QY 185 CAGCTGGCCCTGAACCTGGGCATGGCCCCCCCACATCTTCTTTGGTGCCCCGACAAATGGGCTG 244

Db 387 GCCTTGGCCGATCTTCTGCTTGTCACTGGGTGGGCGGAGCTGCCACAACCCCTGGCTTGGCAC 446

QY 245 GCCTTGGTGATCTCCTGCTGTGTCAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 304

Db 447 CGAGTGAGCCCGCCCGCTGCCCGCTTGCTGTACCCCTTACCTGGCCTGGCTGGCTTTTGGC 506

QY 305 CAGGTGAGCCCGCTGGCGCCCGCCTGCTCTACCCCTACCTGGCCTGGCTGGCCTTCACG 364

Db 507 ACCGTGCTCAACTACTATGTATGTCGCTGCCCGCTTGCTGTACCCCTTACCTGGCCTGGCTGGCCTC 566

QY 365 ACCACACTCAACTACTGCGTATGCGGGGACAACCATGGCTGGGTGGGGGACGGCGGCTG 424

Db 567 GCAGAGTGAAGGCCACCCAGCCATCAGGAATGCAGCCCTGCCAGC 610

QY 425 CCAGAGTGAGTGCCCGGCCACCAGGAGCTGCAGCTGCACCAGC 468

RESULT 11

LOCUS HSPBR3 645 bp DNA PRI 20-MAY-1994

DEFINITION Human peripheral benzodiazepine receptor gene, exon 3.

ACCESSION L21953

NID g483404

KEYWORDS benzodiazepine receptor; peripheral benzodiazepine receptor.

SEGMENT 3 of 4

SOURCE Homo sapiens DNA.

ORGANISM Homo sapiens

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 645)

AUTHORS Lin,D., Chang,Y.J., Strauss,J.F. and Miller,W.L.

TITLE The human peripheral benzodiazepine receptor gene: cloning and characterization of alternative splicing in normal tissues and in a patient with congenital lipoid adrenal hyperplasia

JOURNAL Genomics 18, 643-650 (1993)

MEDLINE 94140364

FEATURES

Location/Qualifiers

1. .645

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/tissue\_type="placenta"

/tissue\_lib="placenta"

224. .362

exon

/product="peripheral benzodiazepine receptor"

BASE COUNT 138 a 177 c 192 g 138 t

ORIGIN

Query Match 19.6%; Score 128; DB 21; Length 645;

Best Local Similarity 98.6%; Pred. No. 8.17e-72;

Matches 140; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Db 223 GGTACGGCTCTACCTGGTCTGGAAGAGCTGGGAGGCTTCACAGAGAAGGCTGTGGTT 282

QY 105 GGTACGGCTCTACCTGGTCTGGAAGAGCTGGGAGGCTTCACAGAGAAGGCTGT-GGTT 163

Db 283 CCCCTGGGCTCTACACTGGGCGAGCTGGCCCTGAACCTGGGCGATGGCCCCCATCTTCTT- 341



QY 164 CCCCTGGGCTCTACACTGGGCAGCTGGCCCTGAACTGGGCATGGGCCATCTTCTTT 223

Db 342 GGTGCCCCGACAAATGGGCTGGG 363  
|||||

QY 224 GGTGCCCCGACAAATGGGCTGGG 245

RESULT 12  
LOCUS RATPTBZR02 3434 bp DNA ROD 25-JAN-1993  
DEFINITION Rat peripheral-type benzodiazepine receptor gene exons 2-4,  
complete cds.  
ACCESSION M84221  
NID 9206478  
KEYWORDS peripheral-type benzodiazepine receptor.  
SEGMENT 2 of 2  
SOURCE Rattus norvegicus (strain Wistar) Adult DNA.  
ORGANISM Rattus norvegicus  
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;  
Murinae; Rattus.  
REFERENCE 1 (bases 1 to 3434)  
AUTHORS Sprengel,R., Werner,P., Seeburg,P.H., Mukhin,A.G., Santi,R.M.,  
Grayson,D.R., Guidotti,A. and Krueger,K.E.  
TITLE Molecular Cloning and Expression of cDNA Encoding a Peripheral-type  
Benzodiazepine Receptor  
JOURNAL J. Biol. Chem. 264, 20415-20421 (1989)  
MEDLINE 90062173  
REFERENCE 2 (bases 1 to 3434)  
AUTHORS Casolotti,S.O., Pelaia,G., Yakovlev,A.G., Csikos,T., Grayson,D.R.  
and Krueger,K.E.  
TITLE Structure of the rat gene encoding the mitochondrial benzodiazepine  
receptor  
JOURNAL Gene 121, 377-382 (1992)  
MEDLINE 93077057  
COMMENT Sequence M84221 starts with a Bam HI site, leaving a gap of about  
7kb between sequences.  
FEATURES  
source  
Location/Qualifiers  
1. .3434  
/organism="Rattus norvegicus"  
/strain="Wistar"  
/db\_xref="taxon:10116"  
/dev\_stage="adult"  
order(M95864:1273. .1462,1. .1094)  
/number=1  
intron  
prim\_transcript 1. .3350  
exon 1095. .1305  
/note="putative"  
/number=2  
join(1124. .1305,1927. .2065,2925. .3113)  
/codon\_start=1  
/product="peripheral-type benzodiazepine receptor"  
/db\_xref="PID:g206480"  
/translation="MSQSWVPVAVGLTVPSLGGFMGAYFVRGGLRWYASLQKPSWHP  
PRWTLAPIWGTLYSAMGYGSYIIWKELGGFTEAMVPLGYTGQLALNWAWPPIFFGA  
ROMGWALVDILMLVSGVATATTLAHRVSPPAARLLYPYLAWLAFATMLNYYVVRDMSG  
RGGGRLTE"  
1306. .1926  
/number=2  
1927. .2065  
/number=3  
2066. .2924  
/number=3  
2925. .3350  
/number=4  
BASE COUNT 713 a 872 c 991 g 858 t  
ORIGIN  
Query Match 17.6%; Score 115; DB 23; Length 3434;  
Best Local Similarity 90.8%; Pred. No. 4.95e-62;  
Matches 128; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Db 1926 GGTATGGCTCCTACATAATCTGGAAGAGCTGGGAGGTTTCACAGAGGCGTATGGTTC 1985  
|||||

QY 105 GGTACGGCTCCTACCTGGTCTGGAAGAGCTGGAGGCTTCACAGAGAGGCTGTGGTTC 164

Db 1986 CCTTGGGTCTCTACACTGGTCAGCTGGCTCTGAAGTGGGCATGGCCCCCATCTTCTTTG 2045  
|||||

QY 165 CCCTGGGCCTCTACACTGGCAGCTGGCCCTGAAGTGGGCATGGCCCCCATCTTCTTTG 224

Db 2046 GTGCCCGGCAGATGGGCTGGG 2066  
|||||

QY 225 GTGCCCGACAAATGGGCTGGG 245

RESULT 13  
LOCUS HSPBR2 487 bp DNA PRI 20-MAY-1994  
DEFINITION Human peripheral benzodiazepine receptor gene, exon 2.  
ACCESSION L21952  
NID 9483403  
KEYWORDS benzodiazepine receptor; peripheral benzodiazepine receptor.  
SEGMENT 2 of 4  
SOURCE Homo sapiens DNA.  
ORGANISM Homo sapiens  
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 487)  
AUTHORS Lin,D., Chang,Y.J., Strauss,J.F. and Miller,W.L.  
TITLE The human peripheral benzodiazepine receptor gene: cloning and  
characterization of alternative splicing in normal tissues and in a  
patient with congenital lipoid adrenal hyperplasia  
JOURNAL Genomics 18, 643-650 (1993)  
MEDLINE 94140364  
FEATURES  
source  
Location/Qualifiers  
1. .487  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/tissue\_type="placenta"  
/tissue\_lib="placenta"  
166. .376  
/product="peripheral benzodiazepine receptor"  
BASE COUNT 67 a 168 c 156 g 96 t  
ORIGIN  
Query Match 16.6%; Score 108; DB 21; Length 487;  
Best Local Similarity 100.0%; Pred. No. 8.35e-57;  
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 272 CCACGGCAGGGTCTCCGCTGGTACGCCGCTGCAGAGCCCTCGTGGCACCCGCCCA 331  
|||||

QY 1 CCACGGCAGGGTCTCCGCTGGTACGCCGCTGCAGAGCCCTCGTGGCACCCGCCCA 60

Db 332 CTGGGTCTGGGCCCTGTCTGGGCACGCTCTACTCAGCCATGGGGTA 379  
|||||

QY 61 CTGGGTCTGGGCCCTGTCTGGGCACGCTCTACTCAGCCATGGGGTA 108

RESULT 14  
LOCUS I66494 7218 bp DNA PAT 23-DEC-1997  
DEFINITION Sequence 14 from patent US 5670367.  
ACCESSION I66494  
NID 92724471  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
Unclassified.  
REFERENCE 1 (bases 1 to 7218)  
AUTHORS Dörner,F., Scheifflinger,F. and Falkner,F.Gunter.  
TITLE Recombinant fowlpox virus  
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;  
FEATURES  
source  
Location/Qualifiers  
1. .7218  
/organism="unknown"  
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others  
ORIGIN  
Query Match 6.0%; Score 39; DB 18; Length 7218;



**THIS PAGE BLANK (USPTO)**

\*\*\*\*\*  
MREI (TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mpsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Jul 21 15:25:04 1998; MasPar time 108.29 Seconds  
Tabular output not generated. 766.833 Million cell updates/sec

Title: >US-09-047-652A-1  
Description: (1-652) from US09047652A.seq  
Perfect Score: 652  
N.A. Sequence: 1 CCACGGCGAAGGCTCCGCT.....GTTCTTGGACATGGAATTT 652  
Comp: GGTGCGGCTCCAGAGCGGA.....CAAGAACCTTGACCTTAAA

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 176886 seqs, 63680241 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-geneseq31-2

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37

Statistics: Mean 8.518; Variance 5.444; scale 1.565

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description Pred. No.
1	238	36.5	272 20 T22195 Human gene signature 3.02e-134
2	44	6.7	91 9 Q51746 Oligonucleotide probe 2.23e-11
3	41	6.3	204 1 N81164 Base substituted E.co 1.04e-09
4	40	6.1	91 9 Q51746 Oligonucleotide probe 3.68e-09
5	36	5.5	114 12 Q70465 Generic DNA sequence 5.38e-07
6	36	5.5	204 1 N81164 Base substituted E.co 5.38e-07
7	35	5.4	114 12 Q70469 Generic DNA sequence 1.83e-06
8	34	5.2	114 12 Q70472 Generic DNA sequence 6.18e-06
9	34	5.2	114 12 Q70472 Generic DNA sequence 6.18e-06
10	34	5.2	114 12 Q70465 Generic DNA sequence 6.18e-06
11	34	5.2	114 12 Q70467 Generic DNA sequence 6.18e-06
12	34	5.2	114 12 Q70467 Generic DNA sequence 6.18e-06
13	34	5.2	114 12 Q70467 Generic DNA sequence 6.18e-06
14	33	5.1	114 12 Q70469 Generic DNA sequence 2.06e-05

c	15	32	4.9	114 12	Q70466	Generic DNA sequence	6.81e-05
	16	32	4.9	114 12	Q70470	Generic DNA sequence	6.81e-05
c	17	32	4.9	178 31	T76405	Human endothelin-1 an	6.81e-05
	18	32	4.9	178 31	T76405	Human endothelin-1 an	6.81e-05
	19	31	4.8	114 12	Q70470	Generic DNA sequence	2.22e-04
	20	31	4.8	172 32	T76363	Human interleukin 8 a	2.22e-04
	21	30	4.6	114 12	Q70473	Generic DNA sequence	7.17e-04
c	22	30	4.6	114 12	Q70466	Generic DNA sequence	7.17e-04
	23	30	4.6	190 31	T76452	Chymase antisense oli	7.17e-04
c	24	29	4.4	114 12	Q70471	Generic DNA sequence	2.28e-03
	25	29	4.4	114 12	Q70473	Generic DNA sequence	2.28e-03
c	26	29	4.4	168 31	T76270	Human MDNCF antisense	2.28e-03
	27	28	4.3	114 12	Q70471	Generic DNA sequence	7.17e-03
	28	28	4.3	128 31	T76233	Human IL6 antisense o	7.17e-03
c	29	27	4.1	114 12	Q70472	Generic DNA sequence	2.22e-02
	30	27	4.1	172 32	T76363	Human interleukin 8 a	2.22e-02
c	31	27	4.1	565 6	Q35072	HCV envelope region n	2.22e-02
	32	26	4.0	89 31	T76219	Human IL5 antisense o	6.77e-02
	33	26	4.0	125 31	T76423	Human endothelin ETA	6.77e-02
c	34	26	4.0	128 31	T76233	Human IL6 antisense o	6.77e-02
c	35	26	4.0	190 31	T76452	Chymase antisense oli	6.77e-02
c	36	26	4.0	379 31	T76527	Inducible nitric oxid	6.77e-02
	37	25	3.8	162 31	T76307	Human RANTES antisens	2.03e-01
c	38	25	3.8	168 31	T76270	Human MDNCF antisense	2.03e-01
	39	25	3.8	250 31	T76438	Substance P antisense	2.03e-01
c	40	25	3.8	250 31	T76438	Substance P antisense	2.03e-01
	41	24	3.7	70 31	T76184	Human IL4 antisense o	5.96e-01
c	42	24	3.7	74 21	T13613	DC43 TSAR library gen	5.96e-01
	43	24	3.7	130 31	T76152	Human vascular cell a	5.96e-01
	44	24	3.7	147 31	T76294	Human defensin 1 anti	5.96e-01
	45	24	3.7	264 31	T76445	Substance P receptor	5.96e-01

ALIGNMENTS

RESULT 1

ID T22195 standard; cDNA to mRNA; 272 BP.

AC T22195;

DT 27-AUG-1996 (first entry)

DE Human gene signature HUMGS03766.

KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;

KW human; cloning; mapping; non-biased library; diagnosis; detection;

KW cell typing; abnormal cell function; ss.

OS Homo sapiens.

PN WO9514772-A1.

PD 01-JUN-1995.

PF 11-NOV-1994; J01916.

PR 12-NOV-1993; JP-355504.

PA (MATS/) MATSUBARA K.

PA (OKUB/) OKUBO K.

PI Matsubara K, Okubo K;

DR WPI; 95-206931/27.

PT Identifying gene signatures in 3'-directed human cDNA library - e.g.

PT for diagnosis of abnormal cell function, by preparing cDNA that

PT reflects relative abundance of corresp. mRNA in specific human

PT tissues

PS Claim 1; Page 1068; 2245pp; Japanese.

CC A single-stranded DNA (or its complementary strand or the corresp.

CC double-stranded DNA) which comprises one of the 7837 "GS" sequences

CC given in T19001-T26837 and which is able to hybridise to part of

CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)

CC sequences were obtained from 3'-directed cDNA libraries prepared

CC from various human tissues; synthesis of cDNA was initiated from the

CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-

CC untranslated sequence is unique to a particular mRNA species, almost

CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library

CC is constructed so as to reflect accurately the relative abundance of

CC different mRNAs in the particular tissue from which it was derived.

CC The appearance frequency of a given GS in a cDNA library can be

CC determined (esp. using primers and probes derived from the GS

CC sequences) as a means of diagnosing abnormal cell function or for

CC recognising different cell types.

SQ Sequence 272 BP; 43 A; 78 C; 86 G; 51 T;

[illegible]

RESULT 2

ID Q51746 standard; cDNA; 91 BP.

AC Q51746;

DT 31-MAY-1994 (first entry)

DE Oligonucleotide probe MK14-A

KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;

KW ss.

OS Synthetic.

PN EP-571911-A.

PD 01-DEC-1993.

PE 24-MAY-1993; 108325.

PR 26-MAY-1992; US-889651.

PA (BECT ) BECTON DICKINSON CO.

PI Shank DD, Spears PA;

DR WPI; 93-378844/48.

PT New oligo:nucleotide probes specific for Mycobacteria - used for

PT detection and amplification of Mycobacteria nucleic acid in

PT samples

PS Claim 3; Page 14; 23pp; English.

CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK1

CC (Q51735). It hybridized to all spp. of mycobacteria tested, but

CC cross reacted to a few non-mycobacterial spp. The probe may

CC be useful as an initial screen for mycobacterial infection.

CC See also Q51735-45 and Q51747-59.

SQ Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

[illegible]

RESULT	3	
ID	N81164	standard; DNA; 204 BP.
AC	N81164;	
DT	08-NOV-1990	(first entry)
DE	Base substituted E.coli beta-galactosidase alpha-fragment.	
KW	E.coli beta galactosidase alpha-fragment; base substitutions; ss.	
OS	Escherichia coli.	
FH	Key	Location/Qualifiers
FT	misc_feature	19..69
FT		/*tag= a
FT		/function= multiple cloning site
FT	primer_bind	187..204

```

FT      EP-285123-A.
PN      05-MAY-1988.
PD      30-MAR-1988; 105163.
PF      03-APR-1987; US-034819.
PR      (SUSO) SUOMEN SOKERI OY.
PA      Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
PI      WPI; 88-279927/40.
DR      Introducing random point mutations into nucleic acids -
PT      by prepn of single stranded template, annealing a primer, elongation,
PT      misincorporation, completion of molecules and screening.
PS      Disclosure; p; English.
CC      Random point mutations were introduced into the alpha fragment of
CC      E. coli beta-galactosidase. The wild type sequence was obtained as a
CC      single stranded template and an oligonucleotide was hybridised to
CC      it to generate a popn of DNA molecules which terminate at all
CC      possible nucleotide positions within a specified region. The
CC      variable 3' ends generated in this way are used as primers for
CC      reverse transcriptase. Nucleotides are misincorporated by the
CC      transcriptase and the molecules are completed to forms that can be
CC      amplified and then expressed in a suitable host-vector system.
CC      The sequence covers all 176 diffit base substitutions, most of which
CC      occurred singularly in any given mutant.
CC      See also P80575.
SQ      Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

      Query Match      6.3%; Score 41; DB 1; Length 204;
      Best Local Similarity 7.7%; Pred. No. 1.04e-09;
      Matches 8; Conservative 60; Mismatches 35; Indels 1; Gaps

Db      86 ymrththyrmrbnvyrdyn-rsdaaaawccyrrsvkdydcynachhddhyvbbvyn 144
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp      190 CAGCTGCCAGGTAGAGGCCCGCCAGGGGAACCCACAGCCCTTCTGTGTGAAGCCTCCCAGCTC 131
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      145 vhnhnncnccbnnhvchvbnhbnhrnwayvrhdarrrdvhcvcv 188
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp      130 TTTCCAGACCAGGTAGGAGCCGCTACCCCATGGCTGAGTAGAGCG 87
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 4

ID Q51746 standard; cDNA; 91 BP.

AC Q51746;

DT 31-MAY-1994 (first entry)

DE Oligonucleotide probe MK14-A

KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;

KW ss.

OS Synthetic.

PN EP-571911-A.

PD 01-DEC-1993.

PF 24-MAY-1993; 108325.

PR 26-MAY-1992; US-889651.

PA (BECT ) BECTON DICKINSON CO.

PI Shank DD, Spears PA;

DR WPI; 93-37884/48.

PT New oligo:nucleotide probes specific for Mycobacteria - used for detection and amplification of Mycobacteria nucleic acid in samples

PT Claim 3; Page 14; 23pp; English.

PS Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14

CC (Q51735). It hybridized to all spp. of mycobacteria tested, but

CC cross reacted to a few non-mycobacterial spp. The probe may

CC be useful as an initial screen for mycobacterial infection.

CC See also Q51735-45 and Q51747-59.

CC Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

SO

	Query Match	6.1%;	Score 40;	DB 9;	Length 91;
	Best Local Similarity	10.7%;	pred. No. 3.68e-09;		
	Matches	6;	Conservative	42;	Mismatches 8; Indels 0; Gaps 0;
Db	10 gssvhsyyvvhhsvshhsvhhvvhhvvhsvvvvhhvvhvhhvvhhvvhsvctcaa	65			
ov	320 gcgcgccgctgtcctacccctacctggccctggctggcttgcttcaggaccactcaaa	375			



```
RESULT 5
ID Q70465 standard; DNA; 114 BP.
AC Q70465;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR petide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PSDB; R65150 and R65151.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PT Disclosure; Page 35; 255pp; English.
CC Q70465 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides.This generic formula can also be
CC represented as follows: X(NNB)6(TGC)(NNB)11Z(NNB)14(TGC)(NNB)3Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active.They may further comprise a linker
CC peptide between the 2 domains.The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed
CC activity allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 5.5%; Score 36; DB 12; Length 114;
Best Local Similarity 4.5%; Pred. No. 5.38e-07;
Matches 5; Conservative 33; Mismatches 74; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbtgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnn 62
: : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 151 CTCTGTGAAGCCTCCAGCTCTTTCCAGACCAGGTAGGAGCCGCTACCCCATGGCTGAGTA 92
: : : : : : : : : : : : : : : : : : : : : : : : : :

Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 114
: : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 91 GAGCGTGCCCCAGACAGGGGCCAGCCAGTGGGGCGGGTGCCACGAGGGC 40

RESULT 6
ID N81164 standard; DNA; 204 BP.
AC N81164;
DT 08-NOV-1990 (first entry)
DE Base substituted E.coli beta-galactosidase alpha-fragment.
KW E.coli beta galactosidase alpha-fragment; base substitutions; ss.
```

```
OS Escherichia coli.
FH Key Location/Qualifiers
FT misc_feature 19..69
FT /*tag= a
FT /function= multiple cloning site
FT primer_bind 187..204
FT /*tag= b
PN EP-285123-A.
PD 05-MAY-1988.
PF 30-MAR-1988; 105163.
PR 03-APR-1987; US-034819.
PA (SUSO) SUOMEN SOKERI OY.
PI Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
DR WPI; 88-279927/40.
PT Introducing random point mutations into nucleic acids -
PT by prepn of single stranded template, annealing a primer, elongation,
PT misincorporation, completion of molecules and screening.
PS Disclosure; p; English.
CC Random point mutations were introduced into the alpha fragment of
CC E.coli beta-galactosidase. The wild type sequence was obtained as a
CC single stranded template and an oligonucleotide was hybridised to
CC it to generate a popn of DNA molecules which terminate at all
CC possible nucleotide positions within a specified region. The
CC variable 3' ends generated in this way are used as primers for
CC reverse transcriptase. Nucleotides are misincorporated by the
CC transcriptase and the molecules are completed to forms that can be
CC amplified and then expressed in a suitable host-vector system.
CC The sequence covers all 176 diffit base substitutions, most of which
CC occurred singularly in any given mutant.
CC See also P80575.
SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 5.5%; Score 36; DB 1; Length 204;
Best Local Similarity 15.1%; Pred. No. 5.38e-07;
Matches 16; Conservative 53; Mismatches 35; Indels 2; Gaps 2;

Db 93 hyrrmrbnvyrdyn-rsdaaawccyrrsvkydcccynachhddhyv-ybbbvynvnhnn 150
: : : : : : : : : : : : : : : : : : : : : : : : : :
QY 19 CTGGTACGGCGGCTGCAGAGCCCTCGTGGCACCCGCCCTGCTGGTGGTGGCCTGT 78
: : : : : : : : : : : : : : : : : : : : : : : : : :

Db 151 cncccbnnhvhnbnnhrnwayvrhdarrddvhccvchccgac 196
: : : : : : : : : : : : : : : : : : : : : : : : : :
QY 79 CTGGGGCAGCTCTACTACGCCATGGGTACGGTCTCCTACCTGGTC 124

RESULT 7
ID Q70469 standard; DNA; 114 BP.
AC Q70469;
DT 07-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6,9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PT Disclosure; Page 35; 255pp; English.
```

CC Q70469 is a generic DNA sequence used to generate random TSAR peptide  
CC This generic formula can be represented as follows: X(TGC)(NNB)10-  
CC (TGC)(NNB)6Z(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction  
CC sites (X is not the same as Y) that are not specified further. This  
CC sequence generates peptides that are cloverleaf in structure. Other  
CC generic sequences are shown in Q70465-68. Other specific peptides  
CC generated by these generic sequences are shown in R65150-54. TSARs are  
CC concatenated heterofunctional proteins or peptides, comprising at least  
CC two functional regions - a binding domain with affinity for a ligand and  
CC a second effector peptide portion that is chemically or biologically  
CC active. They may further comprise a linker peptide between the 2 domains.  
CC The oligonucleotides are also designed so that the expressed peptide  
CC contains 2 or 4 cysteine residues positioned in, or flanking, the  
CC unpredicted or variant residues. These residues confer some degree of  
CC conformational rigidity to the peptides. The TSARs or compsns. comprising  
CC a TSAR binding domain can be used in vivo to deliver a chemically or  
CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin  
CC or enzyme, to the specific target or on the cell. They can also replace  
CC the function of macromolecules, eg. monoclonal or polyclonal antibodies  
CC and therefore circumvent the need for complex methods of hybridoma  
CC formation or in vivo antibody production. The TSARs are easily  
CC characterised and have designed activity allowing direct and rapid  
CC detection in a screening process.  
SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

Query Match 5.4%; Score 35; DB 12; Length 114;  
Best Local Similarity 6.3%; Pred. No. 1.83e-06;  
Matches 7; Conservative 31; Mismatches 73; Indels 0; Gaps 0;

Db 3 cnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62  
QY 16 CCGCTGGTACGCGGCTCTGAGAGCCCTCGTGGCACCGCCGCTGGGTGCTGGGCC 75

Db 63 bnnbtgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbtg 113  
QY 76 TGTCTGGGCGACGCTCTACTCAGCCATGGGGTACGGCTCCTACCTGGTCTG 126

RESULT 8  
ID Q70468 standard; DNA; 114 BP.  
AC Q70468;  
DT 05-APR-1995 (first entry)  
DE Generic DNA sequence to generate a random TSAR peptide library.  
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;  
KW effector domain; concatenated heterofunctional protein; linker;  
KW direct; rapid; detection; screening; treatment; generic; ss.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT misc\_feature 55..60  
FT /\*tag= a  
FT /note= "this sequence represents 'Z'; Z can be a  
FT sequence of 6, 9 or 12 nucleotides (see  
FT comments)"

WO9418318-A.  
18-AUG-1994.  
01-FEB-1994; U00977.  
01-FEB-1993; US-013416.  
30-DEC-1993; US-176500.  
31-JAN-1994; US-189331.  
PA (UYN-) UNIV NORTH CAROLINA.  
PI Fowlkes DM, Kay BK;  
DR WPI; 94-279739/34.  
DR P-PSDB; R65154.  
PN Identifying proteins or peptide(s) which bind a ligand - by  
PD screening a recombinant vector library expressing fusion proteins  
PF comprising a binding domain and an effector domain  
PR Disclosure; Page 35; 255pp; English.  
PR Q70468 is a generic DNA sequence used to generate random TSAR (Totally  
CC Synthetic Affinity Reagents) peptides. This generic formula can also be  
CC represented as follows: X(NNB)11(TGC)(NNB)6Z(NNB)7(TGC)(NNB)10Y. X  
CC and Y are flanking restriction sites (X is not the same as Y) that are  
CC not specified further. Other generic sequences are shown in Q70466-68.  
CC Other specific peptides generated by these generic sequences are shown in

CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,  
CC comprising at least two functional regions - a binding domain with  
CC affinity for a ligand and a second effector peptide portion that is  
CC chemically or biologically active. They may further comprise a linker  
CC peptide between the 2 domains. The oligonucleotides are also designed so  
CC that the expressed peptide contains 2 or 4 cysteine residues positioned  
CC in, or flanking, the unpredicted or variant residues. These residues  
CC confer some degree of conformational rigidity to the peptides. The TSARs  
CC or compsns. comprising a TSAR binding domain can be used in vivo to  
CC deliver a chemically or biologically active moiety, eg. metal ion,  
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the  
CC cell. They can also replace the function of macromolecules, eg.  
CC monoclonal or polyclonal antibodies and therefore circumvent the need  
CC for complex methods of hybridoma formation or in vivo antibody  
CC production. The TSARs are easily characterised and have designed activity  
CC allowing direct and rapid detection in a screening process.  
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 5.2%; Score 34; DB 12; Length 114;  
Best Local Similarity 2.7%; Pred. No. 6.18e-06;  
Matches 3; Conservative 34; Mismatches 75; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnn 62  
QY 16 CCGCTGGTACGCGGCTCTGAGAGCCCTCGTGGCACCGCCGCTGGGTGCTGGGCC 75

Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnb 114  
QY 76 TGTCTGGGCGACGCTCTACTCAGCCATGGGGTACGGCTCCTACCTGGTCTGG 127

RESULT 9  
ID Q70472 standard; DNA; 114 BP.  
AC Q70472;  
DT 10-APR-1995 (first entry)  
DE Generic DNA sequence to generate a random TSAR peptide library.  
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;  
KW effector domain; concatenated heterofunctional protein; linker;  
KW direct; rapid; detection; screening; treatment; generic; ss.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT misc\_feature 55..60  
FT /\*tag= a  
FT /note= "encoded by 2"

WO9418318-A.  
18-AUG-1994.  
01-FEB-1994; U00977.  
01-FEB-1993; US-013416.  
30-DEC-1993; US-176500.  
31-JAN-1994; US-189331.  
PA (UYN-) UNIV NORTH CAROLINA.  
PI Fowlkes DM, Kay BK;  
DR WPI; 94-279739/34.  
DR P-PSDB; R58383.  
PN Identifying proteins or peptide(s) which bind a ligand - by  
PD screening a recombinant vector library expressing fusion proteins  
PF comprising a binding domain and an effector domain  
PR Disclosure; Page 36; 255pp; English.  
PR Q70472 is a generic DNA sequence used to generate random TSAR (Totally  
CC Synthetic Affinity Reagents) peptides. This generic formula can also be  
CC represented as follows: X(NNB)11(CAC)(NNB)11(CAC)(NNB)(CAC)(NNB)22(NNB)6  
CC -(CAC)(NNB)5(CAC)2(NNB)4. X and Y are flanking restriction sites  
CC (X is not the same as Y) that are not specified further. The peptides  
CC generated by this and other generic sequences (Q70470-73) have invariant  
CC histidine residues incorporated into variant sequences. TSARs are  
CC concatenated heterofunctional proteins or peptides, comprising at least  
CC two functional regions - a binding domain with affinity for a ligand and  
CC a second effector peptide portion that is chemically or biologically  
CC active. They may further comprise a linker peptide between the 2 domains.  
CC The TSARs or compsns. comprising a TSAR binding domain can be used in  
CC vivo to deliver a chemically or biologically active moiety, eg. metal  
CC ion, radioisotope, peptide, toxin or enzyme, to the specific target or  
CC on the cell. They can also replace the function of macromolecules, eg.









```

FH Key misc_feature Location/Qualifiers
FT 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; 2 can be a
FT sequence of 6,9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PSDB; R65152.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PT Disclosure; Page 35; 255pp; English.
PS Q70469 is a generic DNA sequence used to generate random TSAR peptide
CC This generic formula can be represented as follows: X(TGC)(NNB)10-
CC (TGC)(NNB)62(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
CC sites (X is not the same as Y) that are not specified further. This
CC sequence generates peptides that are cloverleaf in structure. Other
CC generic sequences are shown in Q70465-68. Other specific peptides
CC generated by these generic sequences are shown in R65150-54. TSARs are
CC concatenated heterofunctional proteins or peptides, comprising at least
CC two functional regions - a binding domain with affinity for a ligand and
CC a second effector peptide portion that is chemically or biologically
CC active.They may further comprise a linker peptide between the 2 domains.
CC The oligonucleotides are also designed so that the expressed peptide
CC contains 2 or 4 cysteine residues positioned in, or flanking, the
CC unpredicted or variant residues. These residues confer some degree of
CC conformational rigidity to the peptides. The TSARs or compens. comprising
CC a TSAR binding domain can be used in vivo to deliver a chemically or
CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC or enzyme, to the specific target or on the cell. They can also replace
CC the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC and therefore circumvent the need for complex methods of hybridoma
CC formation or in vivo antibody production. The TSARs are easily
CC characterised and have designed activity allowing direct and rapid
CC detection in a screening process.
SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

Query Match 5.18; Score 33; DB 12; Length 114;
Best Local Similarity 5.6%; Pred. No. 2.06e-05;
Matches 6; Conservative 30; Mismatches 71; Indels 0; Gaps 0;

Db 6 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnn 65
: : : : : : : : : : : : : : : : : : : : : : : : :
CP 214 GGGGGCCATGCCAGTTCAGGGCCAGCTGCCAGTGTAGAGGCCAGGGGAACACAGC 155
: : : : : : : : : : : : : : : : : : : : : : : : :
Db 66 btgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnt 112
: : : : : : : : : : : : : : : : : : : : : : : : :
CP 154 CTTCTGTGAAGCCTCCAGCTCTTTCCAGACCAGGTAGGAGCCGT 108

RESULT 15
ID Q70466 standard; DNA; 114 BP.
AC Q70466;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR-9 petide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; 2 can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
```

```

PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PSDB; R65152.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PT Disclosure; Page 35; 255pp; English.
PS Q70466 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides.This generic formula can also be
CC represented as follows:X(NNB)1(TGC)(NNB)10(TGC)2(NNB)8(TGC)(NNB)
CC -9Y. X and Y are flanking restriction sites (X is not the same as Y)
CC that are not specified further. Other generic sequences are shown in
CC Q70466-68. Other specific peptides generated by these generic sequences
CC are shown in R65151-54. TSARs are concatenated heterofunctional proteins
CC or peptides, comprising at least two functional regions - a binding
CC domain with affinity for a ligand and a second effector peptide portion
CC that is chemically or biologically active.They may further comprise a
CC linker peptide between the 2 domains.The oligonucleotides are also
CC designed so that the expressed peptide contains 2 or 4 cysteine residues
CC positioned in, or flanking, the unpredicted or variant residues. These
CC residues confer some degree of conformational rigidity to the peptides.
CC The TSARs or compens. comprising a TSAR binding domain can be used in
CC vivo to deliver a chemically or biologically active moiety, eg. metal
CC ion, radioisotope, peptide, toxin or enzyme, to the specific target or
CC on the cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need for
CC complex methods of hybridoma formation or in vivo antibody production.
CC The TSARs are easily characterised and have designed activity allowing
CC direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

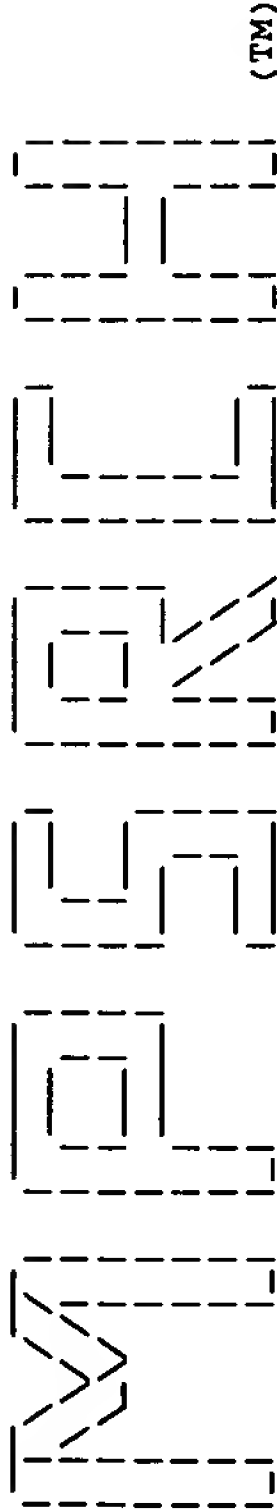
Query Match 4.9%; Score 32; DB 12; Length 114;
Best Local Similarity 8.2%; Pred. No. 6.81e-05;
Matches 9; Conservative 28; Mismatches 73; Indels 0; Gaps 0;

Db 5 gcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnn 64
: : : : : : : : : : : : : : : : : : : : : : : : :
QY 191 GCCCTGAAGTGGGCATGGCCCCCCCCCATCTTCTTGGTCCCCGACAAATGGGCTGGGCCTTG 250
: : : : : : : : : : : : : : : : : : : : : : : : :
Db 65 nbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 114
: : : : : : : : : : : : : : : : : : : : : : : : :
QY 251 GTGGATCTCCTGCTCAGTGGGGGGGGGGGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 300

Search completed: Tue Jul 21 15:26:57 1998
Job time : 113 secs.
```

**THIS PAGE BLANK (USPTO)**

\*\*\*\*\*



\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Jul 21 15:27:14 1998; MasPar time 36.70 Seconds  
Tabular output not generated. 953.206 Million cell updates/sec

Title: >US-09-047-652A-1  
Description: (1-652) from US09047652A.seq  
Perfect Score: 652  
N.A. Sequence: 1 CCACGGCGAAGGTCCTCCGCT.....GTTCTTGGAAACATGGAATTT 652  
Comp: GGTGCCGCTTCCAGAGCGGA.....CAAGAACCTGTACCTTAA

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 104157 seqs, 26825796 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-issued  
1:5\_COMB 2:PCT9\_COMB 3:backfiles

Statistics: Mean 8.113; Variance 4.783; scale 1.696

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	39	6.0	7218	1	US-08-232- Sequence 14, Applicati	2.45e-10
2	36	5.5	215	1	US-08-238- Sequence 5, Applicatio	1.46e-08
3	35	5.4	215	1	US-08-238- Sequence 5, Applicatio	5.58e-08
4	25	3.8	74	2	PCT-US95-1 Sequence 94, Applicati	1.97e-02
5	25	3.8	81	2	PCT-US95-1 Sequence 92, Applicati	1.97e-02
6	24	3.7	74	2	PCT-US95-1 Sequence 94, Applicati	6.50e-02
7	24	3.7	74	2	PCT-US95-1 Sequence 100, Applicat	6.50e-02
8	24	3.7	75	2	PCT-US95-1 Sequence 99, Applicati	6.50e-02
9	24	3.7	81	2	PCT-US95-1 Sequence 92, Applicati	6.50e-02
10	24	3.7	81	2	PCT-US95-1 Sequence 98, Applicati	6.50e-02
11	24	3.7	82	2	PCT-US95-1 Sequence 97, Applicati	6.50e-02
12	23	3.5	66	1	US-08-471- Sequence 144, Applicat	2.10e-01
13	23	3.5	68	1	US-08-471- Sequence 143, Applicat	2.10e-01
14	23	3.5	69	1	US-08-471- Sequence 142, Applicat	2.10e-01
15	23	3.5	74	2	PCT-US95-1 Sequence 100, Applicat	2.10e-01
16	23	3.5	75	2	PCT-US95-1 Sequence 99, Applicati	2.10e-01
17	23	3.5	81	2	PCT-US95-1 Sequence 98, Applicati	2.10e-01
18	23	3.5	82	2	PCT-US95-1 Sequence 97, Applicati	2.10e-01
19	22	3.4	65	1	US-08-471- Sequence 145, Applicat	6.61e-01
20	22	3.4	65	1	US-08-471- Sequence 144, Applicat	6.61e-01

C	21	22	3.4	66	1	US-08-471- Sequence 144, Applicat	6.61e-01
	22	22	3.4	68	1	US-08-471- Sequence 143, Applicat	6.61e-01
	23	22	3.4	69	1	US-08-471- Sequence 142, Applicat	6.61e-01
	24	22	3.4	242	1	US-08-273- Sequence 1, Applicatio	6.61e-01
	25	22	3.4	1995	1	US-08-425- Sequence 3, Applicatio	6.61e-01
C	26	22	3.4	2256	1	US-08-001- Sequence 1, Applicatio	6.61e-01
C	27	22	3.4	2256	1	US-07-794- Sequence 1, Applicatio	6.61e-01
	28	21	3.2	59	2	PCT-US95-1 Sequence 95, Applicati	2.04e+00
C	29	21	3.2	59	2	PCT-US95-1 Sequence 95, Applicati	2.04e+00
	30	21	3.2	66	2	PCT-US95-1 Sequence 93, Applicati	2.04e+00
C	31	21	3.2	66	2	PCT-US95-1 Sequence 93, Applicati	2.04e+00
C	32	21	3.2	108	1	US-08-357- Sequence 2, Applicatio	2.04e+00
C	33	21	3.2	242	1	US-08-273- Sequence 1, Applicatio	2.04e+00
C	34	21	3.2	336	1	US-08-248- Sequence 105, Applicat	2.04e+00
C	35	21	3.2	1441	1	US-08-136- Sequence 18, Applicati	2.04e+00
C	36	21	3.2	1838	2	PCT-US93-0 Sequence 85, Applicati	2.04e+00
	37	21	3.2	2750	1	US-08-136- Sequence 1, Applicatio	2.04e+00
C	38	21	3.2	4175	1	US-08-306- Sequence 49, Applicati	2.04e+00
C	39	21	3.2	4175	2	PCT-US93-0 Sequence 84, Applicati	2.04e+00
C	40	21	3.2	6063	1	US-08-195- Sequence 4, Applicatio	2.04e+00
	41	21	3.2	11558	2	PCT-US93-0 Sequence 23, Applicati	2.04e+00
	42	20	3.1	84	1	US-08-133- Sequence 120, Applicat	6.11e+00
	43	20	3.1	84	1	US-08-300- Sequence 25, Applicati	6.11e+00
C	44	20	3.1	84	2	PCT-US95-1 Sequence 25, Applicati	6.11e+00
C	45	20	3.1	5399	1	US-08-368- Sequence 9, Applicatio	6.11e+00

ALIGNMENTS

RESULT 1  
ID US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.  
AC xxxxxx  
DT  
DE Sequence 14, Application US/08232463  
CC Sequence 14, Application US/08232463  
CC Patent No. 5670367  
CC GENERAL INFORMATION:  
CC APPLICANT: DORNER, F.  
CC APPLICANT: SCHEIFLINGER, F.  
CC APPLICANT: FALKNER, F. G.  
CC TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
CC NUMBER OF SEQUENCES: 52  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Foley & Lardner  
CC STREET: 1800 Diagonal Road, Suite 500  
CC CITY: Alexandria  
CC STATE: VA  
CC COUNTRY: USA  
CC ZIP: 22313-0299  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/232,463  
CC FILING DATE:  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US/07/935,313  
CC FILING DATE:  
CC APPLICATION NUMBER: EP 91 114 300.6  
CC FILING DATE: 26-AUG-1991  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: BENT, Stephen A.  
CC REGISTRATION NUMBER: 29,768  
CC REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (703)836-9300  
CC TELEFAX: (703)683-4109  
CC TELEX: 899149  
CC INFORMATION FOR SEQ ID NO: 14:  
CC SEQUENCE CHARACTERISTICS:

```
CC LENGTH: 7218 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC CLONE: ptzgpt-F1s
SQ SEQUENCE 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 OTHER.

Query Match          6.0%; Score 39; DB 1; Length 7218;
Best Local Similarity 0.3%; Pred. No. 2.45e-10;
Matches 1; Conservative 203; Mismatches 165; Indels 0; Gaps 0;

Db 1067 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1126
QY 13 TCTCCGCTGGTAGCGCGCCTGCAGAGCCCTCGTGGCACCCGCCACTGGGTGCTGGG 72

Db 1127 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1186
QY 73 CCCTGTCTGGGCACGCTCTACTCAGCCATGGGGTACGGCTCCTACCTGGTCTGGAAGA 132

Db 1187 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1246
QY 133 GCTGGGAGGCTTCACAGAGAAGGCTGTGTTCCCTGGCCTCTACACTGGCAGTGCG 192

Db 1247 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1306
QY 193 CCTGAACCTGGGCATGGCCCCCATCTTCTTGGTGCCCGACAAATGGGCTGGCCTTGGT 252

Db 1307 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1366
QY 253 GGATCCTCCTGCTGAGTGGGGCGGGCGGCGGACCCACTACCGTGGCCTGGTACCAGTGAG 312

Db 1367 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1426
QY 313 CCCGCTGGCGCCCGCTGCTCTACCCCTACCTGGCCTGGCTGGCCTTCACGACCACACT 372

Db 1427 YYYYYYYG 1435
QY 373 CAACTACTG 381

RESULT 2
ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
AC xxxxxx
DT
DE Sequence 5, Application US/08238163
CC Sequence 5, Application US/08238163
CC Patent No. 5569830
CC GENERAL INFORMATION:
CC APPLICANT: BENNETT, Alan
CC APPLICANT: LABAVITCH, John M.
CC APPLICANT: POWELL, Ann
CC APPLICANT: STOTZ, Henrik
CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
CC TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
CC NUMBER OF SEQUENCES: 24
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend Khourie and Crew
CC STREET: Steuart Street Tower, One Market Plaza
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: US
CC ZIP: 94105-1493
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/238,163
CC FILING DATE: 03-MAY-1994
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
```

```
CC NAME: Bastian, Kevin L.
CC REGISTRATION NUMBER: 34,774
CC REFERENCE/DOCKET NUMBER: 2307E-540
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 543-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 215 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: misc_feature
CC LOCATION: 1..215
CC OTHER INFORMATION: /standard_name= "Deduced amino acid
CC OTHER INFORMATION: sequence of PGIP from bean."
SQ SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match          5.5%; Score 36; DB 1; Length 215;
Best Local Similarity 13.4%; Pred. No. 1.46e-08;
Matches 28; Conservative 82; Mismatches 98; Indels 1; Gaps 1;

Db 7 SSSVVSRATSCNDKAKKDONTSSWTTDCNRTWGVCDTDTTYRVNNDSGHNKYSSANY 66
Cp 504 CGTGACGGCCACCACATCACAAAGCGTGATGGCACCTGCTGGTGACAGCTCCCTGGT 445

Db 67 NYGNNVGAATHYTHNVSGADSKTVTDSYNASGTSSSNGGTDGNRRSGADSY-GSSKT 125
Cp 444 GGGCGGGGCACTCACTCTGGCAGCGCCGCTCCCCACGCCAGCCATGGTTGTCGCCCAT 385

Db 126 AMTSRNTGKTANNAVDNRNMGDASVSDKNTKKHAKNSADGKVGSKNNGDRNRYGTGT 185
Cp 384 ACGCAGTAGTTGAGTGTGTCGTCTGAAGGCCAGCCAGCCAGGCTAGGGGTAGAGCAGCGG 325

Db 186 KSNVSNCGGKRDVSSYANNKCCGSSC 214
Cp 324 GCGGCCAGCGGGCTCACCTGGTACCAGGC 296

RESULT 3
ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
AC xxxxxx
DT
DE Sequence 5, Application US/08238163
CC Sequence 5, Application US/08238163
CC Patent No. 5569830
CC GENERAL INFORMATION:
CC APPLICANT: BENNETT, Alan
CC APPLICANT: LABAVITCH, John M.
CC APPLICANT: POWELL, Ann
CC APPLICANT: STOTZ, Henrik
CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
CC TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
CC NUMBER OF SEQUENCES: 24
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend Khourie and Crew
CC STREET: Steuart Street Tower, One Market Plaza
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: US
CC ZIP: 94105-1493
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/238,163
CC FILING DATE: 03-MAY-1994
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
```



CC NAME: Bastian, Kevin L.  
CC REGISTRATION NUMBER: 34,774  
CC REFERENCE/DOCKET NUMBER: 2307E-540  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (415) 543-9600  
CC TELEFAX: (415) 543-5043  
CC INFORMATION FOR SEQ ID NO: 5:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 215 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: unknown  
CC MOLECULE TYPE: protein  
CC FEATURE:  
CC NAME/KEY: misc\_feature  
CC LOCATION: 1..215  
CC OTHER INFORMATION: /standard\_name= "Deduced amino acid  
CC OTHER INFORMATION: sequence of PGIP from bean."  
SQ SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match 5.4%; Score 35; DB 1; Length 215;  
Best Local Similarity 12.4%; Pred. No. 5.58e-08;  
Matches 25; Conservative 81; Mismatches 95; Indels 1; Gaps 1;

Db 9 SSVVSRATASCNDKAKKDGNTTSSWTTDCCNRTWGVCDTDTTYRVNNDSGHNKYSSANYNY 68  
QY 304 CCAGGTGAGCCCGCTGGCCGCGCCTGCTCTACCCCTACCTGGCCTGGCTGGCCCTCAC 363  
Db 69 GGNVGAAKTHYYTHTNVSG-ADSKTVTDSYNASGTSSSNGGTDGNRSGADSYGSSKTAM 127  
QY 364 GACCACACTCAACTACTGCGTATGGCGGGGACACCATGGCTGGCGTGGGGACGGCGGCT 423  
Db 128 TSNRTGKTANNAVDSRNMGDASVGSDDKNTKKHAKNSADGKVGSKNNGDRNNRYGTGTS 187  
QY 424 GCCAGAGTGAGTGCCCGGCCACAGGACTGCAGCTGCACCCAGGAGGTGCCATCACGCT 483  
Db 188 NVSNCGGKRDVSSYANKC 209  
QY 484 TGTGATGTGGTGCGCGTACGC 505

RESULT 4  
ID PCT-US95-11934-94 STANDARD; DNA; UNC; 74 BP.  
AC xxxxxx

DE Sequence 94, Application PC/TUS9511934  
CC Sequence 94, Application PC/TUS9511934  
CC GENERAL INFORMATION:  
CC APPLICANT: Cytogen Corporation  
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From  
CC TITLE OF INVENTION: Peptide Libraries  
CC NUMBER OF SEQUENCES: 103  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Pennie & Edmonds  
CC STREET: 1155 Avenue of the Americas  
CC CITY: New York  
CC STATE: New York  
CC COUNTRY: USA  
CC ZIP: 10036  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US95/11934  
CC FILING DATE: 20-SEP-1995  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Misrock, S. Leslie  
CC REGISTRATION NUMBER: 18,872  
CC REFERENCE/DOCKET NUMBER: 1101-196-228  
CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (212) 790-9090  
CC TELEFAX: (212) 869-9741/8864  
CC TELEX: 66141 PENNIE  
CC INFORMATION FOR SEQ ID NO: 94:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 74 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: DNA (genomic)  
SQ SEQUENCE 74 BP; 3 A; 4 C; 3 G; 1 T; 63 OTHER.  
Query Match 3.8%; Score 25; DB 2; Length 74;  
Best Local Similarity 9.9%; Pred. No. 1.97e-02;  
Matches 7; Conservative 20; Mismatches 44; Indels 0; Gaps 0;  
Db 3 GAGNNBN 62  
QY 432 GAGTGCCCGGCCACCCAGGACTGCAGCTGCACCCAGGAGGTGCCATCACGCTGTGTGATGT 491  
Db 63 BNNBNACGCCA 73  
QY 492 GGTGGCCGTCA 502

RESULT 5  
ID PCT-US95-11934-92 STANDARD; DNA; UNC; 81 BP.  
AC xxxxxx  
DE Sequence 92, Application PC/TUS9511934  
CC Sequence 92, Application PC/TUS9511934  
CC GENERAL INFORMATION:  
CC APPLICANT: Cytogen Corporation  
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From  
CC TITLE OF INVENTION: Peptide Libraries  
CC NUMBER OF SEQUENCES: 103  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Pennie & Edmonds  
CC STREET: 1155 Avenue of the Americas  
CC CITY: New York  
CC STATE: New York  
CC COUNTRY: USA  
CC ZIP: 10036  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US95/11934  
CC FILING DATE: 20-SEP-1995  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Misrock, S. Leslie  
CC REGISTRATION NUMBER: 18,872  
CC REFERENCE/DOCKET NUMBER: 1101-196-228  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (212) 790-9090  
CC TELEFAX: (212) 869-9741/8864  
CC TELEX: 66141 PENNIE  
CC INFORMATION FOR SEQ ID NO: 92:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 81 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: DNA (genomic)  
SQ SEQUENCE 81 BP; 3 A; 5 C; 6 G; 4 T; 63 OTHER.

Query Match 3.8%; Score 25; DB 2; Length 81;  
Best Local Similarity 9.9%; Pred. No. 1.97e-02;  
Matches 7; Conservative 20; Mismatches 44; Indels 0; Gaps 0;



```
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 99:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 75 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 75 BP; 1 A; 1 C; 7 G; 5 T; 61 OTHER.

Query Match 3.7%; Score 24; DB 2; Length 75;
Best Local Similarity 6.2%; Pred. No. 6.50e-02;
Matches 4; Conservative 20; Mismatches 41; Indels 0; Gaps 0;

Db 3 GAGNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNN 62
QY 432 GAGTGGCCGGCCACCAGGACTGCAGCTGCACCAGCAGGTCCATCAGCCTTGATGT 491
Db 63 BNBNG 67
QY 492 GGTGG 496

RESULT 9
ID PCT-US95-11934-92 STANDARD; DNA; UNC; 81 BP.
AC xxxxxx
DT
DE Sequence 92, Application PC/TUS9511934
CC Sequence 92, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: CytoGen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 92:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 81 base pairs
CC TYPE: nucleic acid
```

```
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 81 BP; 3 A; 5 C; 6 G; 4 T; 63 OTHER.

Query Match 3.7%; Score 24; DB 2; Length 81;
Best Local Similarity 7.7%; Pred. No. 6.50e-02;
Matches 5; Conservative 20; Mismatches 40; Indels 0; Gaps 0;

Db 16 BNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNA 75
Cp 358 GGGCAGCCAGCCAGGTAGGGGTAGACGAGCGGGCGGCCACCTGGTACCA 299
Db 76 GGCCA 80
Cp 298 GGCCA 294

RESULT 10
ID PCT-US95-11934-98 STANDARD; DNA; UNC; 81 BP.
AC xxxxxx
DT
DE Sequence 98, Application PC/TUS9511934
CC Sequence 98, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: CytoGen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 98:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 81 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 81 BP; 6 A; 6 C; 4 G; 5 T; 60 OTHER.

Query Match 3.7%; Score 24; DB 2; Length 81;
Best Local Similarity 9.1%; Pred. No. 6.50e-02;
Matches 6; Conservative 19; Mismatches 41; Indels 0; Gaps 0;

Db 13 VNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNNN 72
Cp 554 ACCCTGGGCCAAGGCCCTGACAGACTAGCAGGCCAGTGGTCATGAAGCGTGACGGCC 495
Db 73 ACCACA 78
Cp 494 ACCACA 489
```

RESULT 11  
ID PCT-US95-11934-97 STANDARD; DNA; UNC; 82 BP.  
AC xxxxxx  
DT  
DE Sequence 97, Application PC/TUS9511934  
CC Sequence 97, Application PC/TUS9511934  
CC GENERAL INFORMATION:  
CC APPLICANT: Cytogen Corporation  
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From  
CC TITLE OF INVENTION: Peptide Libraries  
CC NUMBER OF SEQUENCES: 103  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Pennie & Edmonds  
CC STREET: 1155 Avenue of the Americas  
CC CITY: New York  
CC STATE: New York  
CC COUNTRY: USA  
CC ZIP: 10036  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US95/11934  
CC FILING DATE: 20-SEP-1995  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Misrock, S. Leslie  
CC REGISTRATION NUMBER: 18,872  
CC REFERENCE/DOCKET NUMBER: 1101-196-228  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (212) 790-9090  
CC TELEFAX: (212) 869-9741/8864  
CC TELEX: 66141 PENNIE  
CC INFORMATION FOR SEQ ID NO: 97:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 82 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: DNA (genomic)  
CC SEQUENCE 82 BP; 1 A; 2 C; 10 G; 8 T; 61 OTHER.  
Query Match 3.7%; Score 24; DB 2; Length 82;  
Best Local Similarity 10.1%; Pred. No. 6.50e-02;  
Matches 7; Conservative 19; Mismatches 43; Indels 0; Gaps 0;  
Db 7 CTCGAGNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNN 66  
QY 43 CTCGTGGCACCCGCCACTGGGTGCTGGGCCCTGTCTGGGGCAGCTCTACTCAGCCAT 102  
Db 67 BNNBNNBGG 75  
QY 103 GGGGTACGG 111  
RESULT 12  
ID US-08-471-052A-144 STANDARD; DNA; UNC; 66 BP.  
AC xxxxxx  
DT  
DE Sequence 144, Application US/08471052A  
CC Sequence 144, Application US/08471052A  
CC Patent No. 5625033  
CC GENERAL INFORMATION:  
CC APPLICANT: Kay, B. K.  
CC APPLICANT: Fowlkes, D. M.  
CC TITLE OF INVENTION: Totally Synthetic Affinity Reagents  
CC NUMBER OF SEQUENCES: 166  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Pennie & Edmonds

CC STREET: 1155 Avenue of the Americas  
CC CITY: New York  
CC STATE: New York  
CC COUNTRY: U.S.A.  
CC ZIP: 10036-2711  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/471,052A  
CC FILING DATE: 06-JUNE-1995  
CC CLASSIFICATION: 530  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Misrock, S. Leslie  
CC REGISTRATION NUMBER: 18,872  
CC REFERENCE/DOCKET NUMBER: 1101-179  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 212 790-9090  
CC TELEFAX: 212 869-8864/9741  
CC TELEX: 66141 PENNIE  
CC INFORMATION FOR SEQ ID NO: 144:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 66 bases  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: unknown  
CC MOLECULE TYPE: DNA  
CC SEQUENCE 66 BP; 2 A; 3 C; 4 G; 2 T; 55 OTHER.  
Query Match 3.5%; Score 23; DB 1; Length 66;  
Best Local Similarity 9.7%; Pred. No. 2.10e-01;  
Matches 6; Conservative 18; Mismatches 38; Indels 0; Gaps 0;  
Db 4 AGNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBCCA 63  
Cp 135 AGCTCTTTCCAGACCAGGTAGGAGCGCGTACCCCATGGCTGAGTAGCGTGCCCCAGACA 76  
Db 64 GG 65  
Cp 75 GG 74  
RESULT 13  
ID US-08-471-052A-143 STANDARD; DNA; UNC; 68 BP.  
AC xxxxxx  
DT  
DE Sequence 143, Application US/08471052A  
CC Sequence 143, Application US/08471052A  
CC Patent No. 5625033  
CC GENERAL INFORMATION:  
CC APPLICANT: Kay, B. K.  
CC APPLICANT: Fowlkes, D. M.  
CC TITLE OF INVENTION: Totally Synthetic Affinity Reagents  
CC NUMBER OF SEQUENCES: 166  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Pennie & Edmonds  
CC STREET: 1155 Avenue of the Americas  
CC CITY: New York  
CC STATE: New York  
CC COUNTRY: U.S.A.  
CC ZIP: 10036-2711  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/471,052A  
CC FILING DATE: 06-JUNE-1995  
CC CLASSIFICATION: 530  
CC ATTORNEY/AGENT INFORMATION:



```
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-179
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212 790-9090
CC TELEFAX: 212 869-8864/9741
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 143:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 68 bases
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: DNA
SQ SEQUENCE 68 BP; 3 A; 3 C; 5 G; 3 T; 54 OTHER.

Query Match 3.5%; Score 23; DB 1; Length 68;
Best Local Similarity 13.4%; Pred. No. 2.10e-01;
Matches 9; Conservative 18; Mismatches 40; Indels 0; Gaps 0;

Db 1 GGTCTAGAVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNV 60
QY 11 GGTCTCCGCTGGTACGGCGGCTGCAGAGCCCTCGTGGCACCGCCCGCCACTGGGTGCTG 70
Db 61 NNACCTG 67
QY 71 GGCCCTG 77

RESULT 14
ID US-08-471-052A-142 STANDARD; DNA; UNC; 69 BP.
AC xxxxxx
DT
DE Sequence 142, Application US/08471052A
CC Sequence 142, Application US/08471052A
CC Patent No. 5625033
CC GENERAL INFORMATION:
CC APPLICANT: Kay, B. K.
CC APPLICANT: Fowlkes, D. M.
CC TITLE OF INVENTION: Totally Synthetic Affinity Reagents
CC NUMBER OF SEQUENCES: 166
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036-2711
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/471,052A
CC FILING DATE: 06-JUNE-1995
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-179
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212 790-9090
CC TELEFAX: 212 869-8864/9741
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 142:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 69 bases
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: DNA
SQ SEQUENCE 69 BP; 2 A; 4 C; 6 G; 2 T; 55 OTHER.
```

```
Query Match 3.5%; Score 23; DB 1; Length 69;
Best Local Similarity 13.6%; Pred. No. 2.10e-01;
Matches 9; Conservative 17; Mismatches 40; Indels 0; Gaps 0;

Db 2 GCTCGAGNNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBN 61
Cp 468 GCTGGTGAGCTGCAGTCCCTGGTGGCGCGGCACTCACTCTGGCAGCCCGCTCCCCCA 409
Db 62 NBCCAG 67
Cp 408 CGCCAG 403

RESULT 15
ID PCT-US95-11934-100 STANDARD; DNA; UNC; 74 BP.
AC xxxxxx
DT
DE Sequence 100, Application PC/TUS9511934
CC Sequence 100, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abptides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 100:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 74 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 74 BP; 6 A; 6 C; 1 G; 1 T; 60 OTHER.

Query Match 3.5%; Score 23; DB 2; Length 74;
Best Local Similarity 9.0%; Pred. No. 2.10e-01;
Matches 6; Conservative 19; Mismatches 42; Indels 0; Gaps 0;

Db 4 GAVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNV 63
QY 414 GACGGCGGCTGCCAGAGTGAGTGCCCGGCCACCAGGAGGACTGCAGCTGCACGAGGTG 473
Db 64 NNACCAC 70
QY 474 CCATCAC 480

Search completed: Tue Jul 21 15:27:55 1998
Job time : 41 secs.
```

---

\*\*\*\*\*  
 M R E N E (TM)  
 \*\*\*\*\*  
 \*\*\*\*\*  
 Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

Mpsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm  
 Run on: Tue Jul 21 15:11:50 1998; MasPar time 760.50 Seconds  
 Tabular output not generated. 1145.170 Million cell updates/sec  
 Title: >US-09-047-652A-1  
 Description: (1-652) from US09047652A.seq  
 Perfect Score: 652  
 N.A. Sequence: 1 CCACGGCGAAGGTCTCCGCT.....GTTCTTGAACATGGAATTT 652  
 Comp: GGTGCCGCTTCCAGAGCGGA.....CAAGAACCTTGACCTTAAA

Scoring table: TABLE default  
 Gap 6  
 Nmatch STD : Dbase 0; Query 0

Searched: 1759237 seqs, 667866413 bases x 2  
 Post-processing: Minimum Match 0%  
 Listing first 45 summaries  
 Database: embl-est54  
 1:em\_est1 2:em\_est3  
 Database: genbank-est106  
 3:gb\_est1 4:gb\_est10 5:gb\_est11 6:gb\_est12 7:gb\_est13  
 8:gb\_est14 9:gb\_est15 10:gb\_est16 11:gb\_est17  
 12:gb\_est18 13:gb\_est19 14:gb\_est2 15:gb\_est20  
 16:gb\_est21 17:gb\_est22 18:gb\_est3 19:gb\_est4 20:gb\_est5  
 21:gb\_est6 22:gb\_est7 23:gb\_est8 24:gb\_est9 25:gb\_gss  
 26:gb\_sts

Statistics: Mean 10.452; Variance 2.420; scale 4.319  
 Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match	Length	DB ID	Description
C 1	498	76.4	546	15	AA724486
C 2	493	75.6	572	9	AA587126
C 3	463	71.0	567	13	AA775735
C 4	459	70.4	563	23	AA161033
C 5	458	70.2	610	12	AA069479
C 6	452	69.3	568	10	AA595715
C 7	446	68.4	591	11	AA643068
C 8	442	67.8	556	8	AA210894
C 9	437	67.0	541	11	AA479680
C 10	425	65.2	541	15	AA748434
C 11	424	65.0	506	7	AA456315
C 12	414	63.5	551	5	AA036726
C 13	405	62.1	505	16	AA838533

C 14	399	61.2	543	15	AA805072	ob86c01.s1	NCI_CGAP_GC	0.00e+00
C 15	394	60.4	507	13	AA768741	ob22g01.s1	NCI_CGAP_K1	0.00e+00
C 16	392	60.1	558	11	AA479803	zu43a07.r1	Soares ovar	0.00e+00
C 17	390	59.8	457	7	AA455945	aa16e07.s1	Soares NHM	0.00e+00
C 18	386	59.2	505	16	AA768733	oc86h05.s1	NCI_CGAP_GC	0.00e+00
C 19	386	59.2	516	9	AA595368	no36d03.s1	NCI_CGAP_Pr	0.00e+00
C 20	376	57.7	540	12	AA654418	nt03h01.s1	NCI_CGAP_Ly	0.00e+00
C 21	375	57.5	465	11	AA653746	ns63d09.s1	NCI_CGAP_Pr	0.00e+00
C 22	373	57.2	521	9	AA559042	nl11e02.s1	NCI_CGAP_Pr	0.00e+00
C 23	364	55.8	504	15	AA805330	oc15c05.s1	NCI_CGAP_GC	0.00e+00
C 24	360	55.2	490	15	AA857072	oe33d06.s1	NCI_CGAP_Pr	0.00e+00
C 25	360	55.2	498	16	AA808022	oc38e02.s1	NCI_CGAP_GC	0.00e+00
C 26	356	54.6	481	15	AA834034	of25e01.s1	NCI_CGAP_K1	0.00e+00
C 27	346	53.1	541	9	AA535901	nf94b05.s1	NCI_CGAP_Co	0.00e+00
C 28	345	52.9	444	22	AA100452	zl81c10.s1	Stratagene	0.00e+00
C 29	342	52.5	401	17	AA904236	oe73g05.s1	NCI_CGAP_Lu	0.00e+00
C 30	341	52.3	419	22	AA099285	zl77h01.s1	Stratagene	0.00e+00
C 31	337	51.7	424	23	AA158134	zo43h11.s1	Stratagene	0.00e+00
C 32	320	49.1	495	23	AA158833	zo58e04.r1	Stratagene	0.00e+00
C 33	315	48.3	402	19	N23981	yx71g01.s1	Homo sapien	0.00e+00
C 34	314	48.2	416	22	W44958	zc20g05.s1	Soares sene	0.00e+00
C 35	312	47.9	323	4	AA313263	EST185182	Colon carcin	0.00e+00
C 36	311	47.7	446	16	AA837816	oe39h02.s1	NCI_CGAP_Pr	0.00e+00
C 37	308	47.2	471	22	AA056606	zl66c07.s1	Stratagene	0.00e+00
C 38	308	47.2	541	22	W25523	zb90b12.r1	Soares sene	0.00e+00
C 39	307	47.1	462	20	N63826	za26e09.s1	Homo sapien	0.00e+00
C 40	306	46.9	374	21	W80770	zd90c02.s1	Soares feta	0.00e+00
C 41	306	46.9	578	8	AA210893	zr90g06.r1	NCI_CGAP_GC	0.00e+00
C 42	305	46.8	468	22	AA037186	zc51c06.r1	Soares sene	0.00e+00
C 43	297	45.6	423	22	W70148	zd52g04.s1	Soares feta	0.00e+00
C 44	291	44.6	460	20	N94184	za26e09.r1	Homo sapien	0.00e+00
C 45	286	43.9	334	7	AA526719	ni91c12.s1	NCI_CGAP_Pr	0.00e+00

ALIGNMENTS

RESULT 1  
 LOCUS AA724486 546 bp mRNA EST 11-FEB-1998  
 DEFINITION ah99h11.s1 Soares NFL T GBC S1 Homo sapiens cDNA clone 1327269 3' similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR (HUMAN);, mRNA sequence.  
 ACCESSION AA724486  
 NID g2742193  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 546)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Tel: (301) 496-1550  
 Email: Robert.Strausberg@nih.gov  
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
 Seq primer: -40ml3 fwd. ET from Amersham  
 High quality sequence stop: 472.

FEATURES  
 source  
 Location/Qualifiers  
 1..546  
 /organism="Homo sapiens"  
 /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbHL19W, testis NHT, and B-cell NCI\_CGAP\_GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of

I.M.A.G.E. clones 297480-302087, 682632-687239,  
726408-728711, and 729096-731399. Subtraction by Bento  
Soares and M. Fatima Bonaldo."  
/db\_xref="taxon:9606"  
/clone="1327269"  
/clone\_lib="Soares NFL T GBC S1"  
/lab\_host="DH10B"  
/lab\_host="DH10B"

BASE COUNT 120 a 166 c 166 g 91 t 3 others  
ORIGIN

Query Match 76.4%; Score 498; DB 15; Length 546;  
Best Local Similarity 98.4%; Pred. No. 0.00e+00;  
Matches 503; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Db 36 AAATTCCATGTTCCAAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 95  
|||||  
Cp 652 AAATTCCATGTTCCAAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 593  
Db 96 GTGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCAAGGCCCTGAC 155  
|||||  
Cp 592 GTGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCAAGGCCCTGAC 533  
Db 156 AGACTAGCAGGCCCTCAGTGGTTCATGAAGCGGTGACGGCCACCACATCACAAGCGTGATGGC 215  
|||||  
Cp 532 AGACTAGCAGGCCCTCAGTGGTTCATGAAGCGGTGACGGCCACCACATCACAAGCGTGATGGC 473  
Db 216 ACCTGCTGGTCANGTCAGTCCCTGGTGGGGCGGCGCACTCACTCTGGCAGCCGCGTCC 275  
|||||  
Cp 472 ACCTGCTGGTCANGTCAGTCCCTGGTGGGGCGGCGCACTCACTCTGGCAGCCGCGTCC 413  
Db 276 CCCACGCCAGCCATGGTTGTCCTCCGCCATACGCAGTAGTTGAGTGTGGTCTGTGAAGGCCAG 335  
|||||  
Cp 412 CCCACGCCAGCCATGGTTGTCCTCCGCCATACGCAGTAGTTGAGTGTGGTCTGTGAAGGCCAG 353  
Db 336 CCAGGCCAGGTAGGGGTAGACAGAGCGGGGGCGGCGGCTCACCTGGTACCAGGCCAC 395  
|||||  
Cp 352 CCAGGCCAGGTAGGGGTAGACAGAGCGGGGGCGGCGGCTCACCTGGTACCAGGCCAC 293  
Db 396 GGTAGTGTGNCNCGCCGCCCACTGACACAGCAGGAGATCCACCAAGGCCAGCCCAATTG 455  
|||||  
Cp 292 GGTAGTGGTGTCCCGCGCCCACTGACACAGCAGGAGATCCACCAAGGCCAGCCCAATTG 233  
Db 456 TCGGGCACCAAGAAAGATGGGGGGCCCATGCCAGTTCAGGGCCAGCTGCCAGTGTAGAG 515  
Cp 232 TCGGGCACCAAGAAAGATGGGGGGCCCATGCCAGTTCAGGGCCAGCTGCCAGTGTAGAG 173  
Db 516 GCCCAGGGGAACCCACAGCCTTCTCTGTGAAG 546  
|||||  
Cp 172 GCCCAGGGGAACCCACAGCCTTCTCTGTGAAG 142

RESULT 2  
LOCUS AA587126 572 bp mRNA EST 12-SEP-1997  
DEFINITION nn70b12.s1 NCI\_CGAP\_Lar1 Homo sapiens cDNA clone IMAGE:1089215  
similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR  
(HUMAN);, mRNA sequence.  
ACCESSION AA587126  
NID g2397940  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 572)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT

Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert\_Strausberg@nih.gov  
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.

Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: Stratagene, Inc.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 462.

FEATURES  
source

Location/Qualifiers  
1..572  
/organism="Homo sapiens"  
/note="Organ: larynx; Vector: Bluescript SK-; Site\_1:  
EcORI; Site\_2: XhoI; Cloned unidirectionally. Primer:  
Oligo dt. Larynx. 5' adaptor sequence: 5' GAATTGGGCACGAG  
3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'  
Average insert size: 0.9 kb."  
/db\_xref="taxon:9606"  
/clone="IMAGE:1089215"  
/clone\_lib="NCI\_CGAP\_Lar1"  
/tissue\_type="larynx"  
/lab\_host="SOLR (kanamycin resistant)"  
<1..>572

BASE COUNT 132 a 171 c 173 g 96 t  
ORIGIN

Query Match 75.6%; Score 493; DB 9; Length 572;  
Best Local Similarity 96.5%; Pred. No. 0.00e+00;  
Matches 522; Conservative 0; Mismatches 17; Indels 2; Gaps 2;

Db 34 AAATTCCATGTTCCAAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 93  
|||||  
Cp 652 AAATTCCATGTTCCAAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 593  
Db 94 GTGGGGCTCAGGTGGGGCAACCTCTGAAGCTCTGTGACCCCTGGGCCAAGGCCCTGAC 153  
|||||  
Cp 592 GTGGGGCTCAGGTGGGGCAACCTCTGAAGCTCTGTGACCCCTGGGCCAAGGCCCTGAC 533  
Db 154 AGACTAGCAGGCCCGCAGTGGTCATGAAAGCGTGACGGCCACCACATCACAAGCGTGATGGC 213  
|||||  
Cp 532 AGACTAGCAGGCCCGCAGTGGTCATGAAAGCGTGACGGCCACCACATCACAAGCGTGATGGC 473  
Db 214 ACCTGCTGGTGCAGCTGCAGTCCCTGGTGGGGCGGCGCACTCACTCTGGCAGCCCGCTCC 273  
|||||  
Cp 472 ACCTGCTGGTGCAGCTGCAGTCCCTGGTGGGGCGGCGCACTCACTCTGGCAGCCCGCTCC 413  
Db 274 CCCATGCCAGCCATGGTTGTCCCGCCATACGCAGTAGTGTGAGTGTGGTCGCGAAGGCCAG 333  
|||||  
Cp 412 CCCACGCCAGCCATGGTTGTCCCGCCATACGCAGTAGTGTGAGTGTGGTCGTAAGGCCAG 353  
Db 334 CCAGGCCAGGTAGGGGTAGAGCAGGGGGGGCGGCCAGCGGGCTCACCTGGTACCAGGCCAC 393  
|||||  
Cp 352 CCAGGCCAGGTAGGGGTAGAGCAGGGGGGGCGGCCAGCGGGCTCACCTGGTACCAGGCCAC 293  
Db 394 GGTAGTGTGGACCGCGCCCTCACTGACCGAGGAGATCCACCAAGGCCCGCCATTG 453  
|||||  
Cp 292 GGTAGTGGTGTCCCGCCCGCCCTCACTGACCGAGGAGATCCACCAAGGCCCGCCATTG 233  
Db 454 TCGGGCACCAAAAGAGATGGGGAGGCCATGCCAGTTCAGGGCGAGCTGCACAGTGTAGAG 513  
|||||  
Cp 232 TCGGGCACCAAAAGAGATGGGGGGCCATGCCAGTTCAGGGCGAGCTGCCAGTGTAGAG 173  
Db 514 GCCCAGGGGAAC-ACAGCCTTCTGTGTGAAGACTCCAGCTCTTTC-AGACCAGGTATGA 571  
|||||  
Cp 172 GCCCAGGGGAACACAGCCTTCTGTGTGAAGCCTCCAGCTCTTTCAGACCAGGTATGA 113  
Db 572 G 572  
Cp 112 G 112

RESULT 3



LOCUS AA775735 567 bp mRNA EST 05-FEB-1998  
DEFINITION zf31f04.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone  
378559 3' similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE  
RECEPTOR (HUMAN);, mRNA sequence.  
ACCESSION AA775735  
NID 92835069  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 567)  
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,  
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,  
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,  
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.  
TITLE WashU-NCI human EST Project  
JOURNAL Unpublished (1997)  
COMMENT  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 487.  
FEATURES  
source  
1. .567  
/organism="Homo sapiens"  
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a  
modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer [5'  
TGTTACCAATCTGAAGTGGGAGCGCGCATCTTTTCTTTTCTTTT 3'],  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified pT73 vector  
(Pharmacia). Library went through one round of  
normalization to a Cot = 5. Library constructed by  
M.Fatima Bonaldo. This library was constructed from the  
same fetus as the fetal lung library, Soares fetal lung  
NbHL19W."  
/db\_xref="GDB:1286815"  
/db\_xref="taxon:9606"  
/clone="378559"  
/clone\_lib="Soares fetal heart NbHH19W"  
/sex="unknown"  
/dev\_stage="19 weeks"  
/lab\_host="DH10B (ampicillin resistant)"  
BASE COUNT 126 a 167 c 175 g 98 t 1 others  
ORIGIN  
Query Match 71.0%; Score 463; DB 13; Length 567;  
Best Local Similarity 96.3%; Pred. NO. 0.00e+00;  
Matches 522; Conservative 0; Mismatches 12; Indels 8; Gaps 8;  
Db 30 AAATTCCATGTTCCAGAACATGCTCTAAGCATGCGAGAACACAGGACACTGCTCCCGG 89  
|||||  
Cp 652 AAATTCCATGTTCCAGAACATGCTCTAAGCATGCGAGAACACAGGACACTGCTCCCGG 593  
Db 90 GTGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCGCTGTAC 149  
|||||  
Cp 592 GTGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCGCTGAC 533  
Db 150 AGACTAGCAGGCCCGCAGTGGTGCATGAAGCGGTGACGGCCACCATCATCACAAGCGTGATGGC 209  
|||||  
Cp 532 AGACTAGCAGGCCCGCAGTGGTGCATGAAGCGGTGACGGCCACCATCATCACAAGCGTGATGGC 473  
Db 210 ACCTGCTGGTG-AGCTGCAGTCCCTGGTGGGCCGGGGCACTCACTCTGGCAGCCCGCTCC 268  
|||||  
Cp 472 ACCTGCTGGTGACGCTGCAGTCCCTGGTGGGCCGGGGCACTCACTCTGGCAGCCCGCTCC 413

Db 269 CCCACGCCAGCCATGGTTGTCCCGCCATAGCAGTAGTTAGTGTGCTGAAGGCCAG 328  
|||||  
Cp 412 CCCACGCCAGCCATGGTTGTCCCGCCATAGCAGTAGTTAGTGTGCTGAAGGCCAG 353  
Db 329 CCAGGCTAGGTAGGGTAGAGCAGGCGGGGCCAGCGGGCTCACCTGGTACCAGGCCAC 388  
|||||  
Cp 352 CCAGGCCAGGTAGGGTAGAGCAGGCGGGGCCAGCGGGCTCACCTGGTACCAGGCCAC 293  
Db 389 GGTAGTGTGACTCGCCGCCCCACTGACCAGCAGGAGATCCACCAAGGCCAGCCCATTTG 448  
|||||  
Cp 292 GGTAGTGGCTGCGCGCGCCCTGACCTGACCAGCAGGAGATCCACCAAGGCCAGCCCATTTG 233  
Db 449 TCGGGCAC-ANAGAAGATGGGGGCCCATGCC-AGTTCAGGGGCCAGCTGCCTAGTGTAGAG 506  
|||||  
Cp 232 TCGGGCACCAAGAAGATGGGGGCCCATGCCAGTTCAGGGGCCAGCTGCCCAGTGTAGAG 173  
Db 507 GCC-AGAGGAAC-ACAGCCTTCTCTGTGAAGCCTCC-AGCTCTTTC-AGAC-AGGTAGGA 561  
|||||  
Cp 172 GCCAGGGGAACACAGCCTTCTCTGTGAAGCCTCCAGCTCTTCCAGACCAGGTAGGA 113  
Db 562 GC 563  
||  
Cp 112 GC 111  
RESULT 4  
LOCUS AA161033 563 bp mRNA EST 16-DEC-1996  
DEFINITION z058e04.s1 Stratagene pancreas (#937208) Homo sapiens cDNA clone  
591102 3' similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE  
RECEPTOR (HUMAN);, mRNA sequence.  
ACCESSION AA161033  
NID g1735356  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 563)  
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,  
Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,  
Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.  
TITLE WashU-Merck EST Project  
JOURNAL Unpublished (1995)  
COMMENT  
Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40M13 fwd. from Amersham  
High quality sequence stop: 425.  
FEATURES  
Location/Qualifiers  
1. .563  
/organism="Homo sapiens"  
/note="Organ: pancreas; Vector: pBluescript SK-; Site\_1:  
EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer:  
Oligo dT. Pancreatic adenocarcinoma cell line. Average  
Insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor  
sequence: 5' GAATTCGCGCAGAG 3' -3' adaptor sequence: 5'  
CTCGAGTTTCTTTTCTTTTCTTTT 3'"  
/db\_xref="taxon:9606"  
/clone="591102"  
/clone\_lib="Stratagene pancreas (#937208)"  
/lab\_host="SOLR cells (kanamycin resistant)"  
complement(<1.>563)  
mRNA 127 a 168 c 169 g 98 t 1 others  
BASE COUNT  
ORIGIN

Query Match 70.4%; Score 459; DB 23; Length 563;  
Best Local Similarity 96.2%; Pred. No. 0.00e+00;  
Matches 513; Conservative 0; Mismatches 13; Indels 7; Gaps 7;

Db 35 AAATTCCATGTTCCAGAACATGCTCTAAGCATGCGAAGACACAGGACACTGCTCCCGG 94  
|||||  
Cp 652 AAATTCCATGTTCCAGAACATGCTCTAAGCATGCGAAGACACAGGACACTGCTCCCGG 593  
|||||  
Db 95 GTGGGGGCTCAGGTGGGCAACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCCATGAC 154  
|||||  
Cp 592 GTGGGGGCTCAGGTGGGCGCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCCCTGAC 533  
|||||  
Db 155 AGACTAGCAGGCCCTCAGTGGTCATGAAGCGTGACGGCCACCATCATCAAGCGTGATGGC 214  
|||||  
Cp 532 AGACTAGCAGGCCCTCAGTGGTCATGAAGCGTGACGGCCACCATCATCAAGCGTGATGGC 473  
|||||  
Db 215 ACCTGCTGGTN-AGCTGCAGTCCCTGGTGGGCGGGCAGCTCACTCTGGCAGCGCCGCTCC 273  
|||||  
Cp 472 ACCTGCTGGTGACGCTGCAGTCCCTGGTGGGCGGGCAGCTCACTCTGGCAGCGCCGCTCC 413  
|||||  
Db 274 CCATGCCAGCCATGTTGCCCGCCATACGCAGTAGTGTGTTGGTGGCGAAGGCCAG 333  
|||||  
Cp 412 CCACGCCAGCCATGTTGTCGCCCATACGCAGTAGTGTGTTGGTGGTGAAGGCCAG 353  
|||||  
Db 334 CCAGGCCAGGTAAGGGGTAGCAGGCGGGCGGCCAGCGGGGCTTCACTGTGTACCAGGC 393  
|||||  
Cp 352 CCAGGCCAGGTA-GGGGTAGCAGGCGGGCGGCCAGCGGG-CT-CACCTGGTACCAGGC 296  
|||||  
Db 394 CACGGTAGTTGGC-CCGCCGCCCTACTGACCAGCAGGAGATCCACCAAGGCCAGCCCAT 452  
|||||  
Cp 295 CACGGTAGTTGGTGC CGCGCCCTACTGACCAGCAGGAGATCCACCAAGGCCAGCCCAT 236  
|||||  
Db 453 TTGTCGGGCACCAAGAAGATGGGGGCCATGCAAGTTTCAGGGCCAGTGCC-AGTGTA 511  
|||||  
Cp 235 TTGTCGGGCACCAAGAAGATGGGGGCCATGCCAGTTTCAGGGCCAGTGCCAGTGTA 176  
|||||  
Db 512 GAGGCCAGGGGAACACACAGCCTTCTGTGAAGCTTCC-AGCTCTTTCCAGA 563  
|||||  
Cp 175 GAGGCCAGGGGAACACACAGCCTTCTGTGAAGCCTCCAGCTCTTTCCAGA 123  
|||||

RESULT 5  
LOCUS AA069479 610 bp mRNA EST 23-DEC-1997  
DEFINITION zml4h11.s1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 525669 3' similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR (HUMAN);, mRNA sequence.

ACCESSION AA069479  
NID g1577104  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 610)  
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.  
WashU-Merck EST Project  
Unpublished (1995)

Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
WARNING: There is evidence that suggests that the 384-well parent plate of this clone contains both human and mouse derived clones. Thus, the origin of this clone is uncertain. This caution should be kept in mind should you use this clone.

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 951 Std Error: 0.00  
Seq primer: -40M13 fwd. from Amersham  
High quality sequence stop: 118.  
FEATURES  
Location/Qualifiers  
1..610  
/organism="Homo sapiens"  
/note="Organ: pancreas; Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Pancreatic adenocarcinoma cell line. Average insert size: 1.0 kb; Uni-2AP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTTCTTTTCTTTTCTTTT 3'."  
/db\_xref="GDB:3917098"  
/db\_xref="taxon:9606"  
/clone="525669"  
/clone\_lib="Stratagene pancreas (#937208)"  
/lab\_host="SOLR cells (kanamycin resistant)"  
BASE COUNT 133 a 177 c 175 g 103 t 22 others  
ORIGIN

Query Match 70.2%; Score 458; DB 12; Length 610;  
Best Local Similarity 92.3%; Pred. No. 0.00e+00;  
Matches 527; Conservative 0; Mismatches 35; Indels 9; Gaps 8;

Db 33 AAATTCCATGTTCCAGAACATGCTCTAAGCATGCGAAGACACAGGACACTGCTCCCGG 92  
|||||  
Cp 652 AAATTCCATGTTCCAGAACATGCTCTAAGCATGCGAAGACACAGGACACTGCTCCCGG 593  
|||||  
Db 93 GTGGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCC-TGAC 151  
|||||  
Cp 592 GTGGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCCCTGAC 533  
|||||  
Db 152 AGACTANAGGCCAGTGGTGCATGAAGCGTGACGGCCACCATCATCAAGCGTGATGGC 211  
|||||  
Cp 532 AGACTAGCAGGCCAGTGGTGCATGAAGCGTGACGGCCACCATCATCAAGCGTGATGGC 473  
|||||  
Db 212 ACCTGCTGGT-CAGCTGCAGTCCCTGGTGGCGGGGCACCTCACTCTGGCAGNCGCGTCC 270  
|||||  
Cp 472 ACCTGCTGGTGCAGTGCAGTCCCTGGTGGCGGGGCACCTCACTCTGGCAGNCGCGTCC 413  
|||||  
Db 271 CCCACGCCAGCCATNGTTGTCGCCCTACTAGTAGTTGAGTNTGGTCGGAAGGCCAG 330  
|||||  
Cp 412 CCCACGCCAGCCATGGTTGTCCGCCATACGCAGTAGTTGAGTGTGGTGGTGAAGGCCAG 353  
|||||  
Db 331 CCAGGCCAGGTATGGGTAGACAGGCGGGGCCAGCGGGCTCACCTGGTACCAGGCCAC 390  
|||||  
Cp 352 CCAGGCCAGGTAGGGGTAGACAGGCGGGGCCAGCGGGCTCACCTGGTACCAGGCCAC 293  
|||||  
Db 391 GGTATTGTC--CCGCCGCCCTACCTNACGAGNANNAGGATTCAACCAAGGCCAGCCCATTT 448  
|||||  
Cp 292 GGTAGTGGTGC CGCGCGCCCTGACCTGACCAAGCAGGAG-ATCCACCAAGGCCAGCCCATTT 234  
|||||  
Db 449 GTCGGGGACCAAGAAGATNGGGGGCATGCC-AGTTCAGGGC-AGNTGCCNAGTGTAGA 506  
|||||  
Cp 233 GTCGGGGACCAAGAAGATNGGGGGCATGCCAGTTCAGGGCCAGCTGCCAGTGTAGA 174  
|||||  
Db 507 GGCCCCAGGTGAACACANGCTTCTCTGTGNAGC-TCCCAG-TNTTTCCAGACCAGTNNGG 564  
|||||  
Cp 173 GGCCCCAGGGGAACCAAGCCTTCTCTGTGAAGCCTCCCAGCTCTTTTCCAGACCAGGTAGG 114  
|||||  
Db 565 AGCCGTNNCCATGGCTANGTAGAGCTTCCC 595  
|||||  
Cp 113 AGCCGTACCCCATGGCTGAGTAGAGCGTGCC 83  
|||||

RESULT 6  
LOCUS AA595715 568 bp mRNA EST 25-SEP-1997  
DEFINITION n146e08.s1 NCI\_CGAP\_Lul Homo sapiens cDNA clone IMAGE:979910 similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR (HUMAN);, mRNA sequence.  
ACCESSION AA595715

NID 92411065  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 568)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert\_Strausberg@nih.gov  
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: Stratagene, Inc., David B. Krizman,  
Ph.D.  
CDNA Library Arraying: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Insert Length: 713 Std Error: 0.00  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 265.  
FEATURES  
source  
1..568  
/organism="Homo sapiens"  
/note="Organ: lung; Vector: Bluescript SK-; Site\_1: EcoRI;  
Site\_2: XhoI; Cloned unidirectionally. Primer: Oligo dt.  
Bulk lung tumor. 5' adaptor sequence: 5' GAATTCGGCAGG 3'  
3' adaptor sequence: 5' CTCGAGTTTITTTTTTTTTTTT 3'  
Average insert size: 1.1 kb."  
/db\_xref="taxon:9606"  
/clone="IMAGE:979910"  
/clone\_lib="NCI-CGAP\_Lu1"  
/tissue\_type="lung tumor"  
/lab\_host="SOLR (kanamycin resistant)"  
BASE COUNT 124 a 164 c 165 g 113 t 2 others  
ORIGIN  
Query Match 69.3%; Score 452; DB 10; Length 568;  
Best Local Similarity 96.9%; Pred. No. 0.00e+00;  
Matches 496; Conservative 0; Mismatches 10; Indels 6; Gaps 6;  
Db 53 AAATTCATGTTCCAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 112  
|||||  
Cp 652 AAATTCATGTTCCAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 593  
Db 113 GTGGGGCTCAGGTGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCCCTGAC 172  
|||||  
Cp 592 GTGGGGCTCAGGTGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCCCTGAC 533  
Db 173 AGACTAGCAGGCCCGAGTGGTCATGAAGCGTGACGGCCACCATCATCACAAGCGTGATGGC 232  
|||||  
Cp 532 AGACTAGCAGGCCCGAGTGGTCATGAAGCGTGACGGCCACCATCATCACAAGCGTGATGGC 473  
Db 233 ACCTGCTGGTGTAGCTGCAGTCCCTGGTGGGCGGGGCACTACTCTGGCAGCCCGCTCC 292  
|||||  
Cp 472 ACCTGCTGGTGCAGCTGCAGTCCCTGGTGGGCGGGGCACTACTCTGGCAGCCCGCTCC 413  
Db 293 CCCAGCCAGCCATGGTTGTCGGCCATACGCAGTAGTGTAGTGTGGTGTGAAGGCCAG 352  
|||||  
Cp 412 CCCAGCCAGCCATGGTTGTCGGCCATACGCAGTAGTGTAGTGTGGTGTGAAGGCCAG 353  
Db 353 CCAGGCCAGGTAGGGGTAGAGCAGCGGGCGGCCAGCGGGCTCACCTGGTACCAGGCCAC 412  
|||||  
Cp 352 CCAGGCCAGGTAGGGGTAGAGCAGCGGGCGGCCAGCGGGCTCACCTGGTACCAGGCCAC 293  
Db 413 GGTACTGGNN-CCGCCGCCCTACTGTACCAGCAGGAGATCCACCAAGGCCCATTTG 471

Cp 292 GGTAGTGGCTGCCGCCGCCCTACTGACCAAGGAGATCCACCAAGGCCCATTTG 233  
Db 472 TCGGGCACCAAGAAGATGGGGGCCCATGCAC-GTTCTAGGGCAGTGCCAA-TGTAGAG 529  
|||||  
Cp 232 TCGGGCACCAAGAAGATGGGGGCCCATGCCAGTTCAGGGCCAGCTGCCAGTGTAGAG 173  
Db 530 GCCAAGGG-AAC-ACAGC-TTCTCTGTGAAGC 558  
|||||  
Cp 172 GCCCAGGGGAACACACAGCCTTCTCTGTGAAGC 141  
RESULT 7  
LOCUS AA643068 591 bp mRNA EST 27-OCT-1997  
DEFINITION nr95f04.s1 NCI-CGAP\_Pr25 Homo sapiens cDNA clone IMAGE:1175743  
similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR  
(HUMAN);, mRNA sequence.  
ACCESSION AA643068  
NID g2568286  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 591)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert\_Strausberg@nih.gov  
Tissue Procurement: Suzanne L. Topalian, M.D., Robert K. Bright,  
Ph.D.  
CDNA Library Preparation: Stratagene, Inc.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 337.  
FEATURES  
source  
Location/Qualifiers  
1..591  
/organism="Homo sapiens"  
/note="Organ: prostate; Vector: Bluescript SK-; Site\_1:  
EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer:  
Oligo dt. Normal prostate epithelial cell line (HPV  
immortalized). 5' adaptor sequence: 5' GAATTCGGCAGG 3'  
3' adaptor sequence: 5' CTCGAGTTTITTTTTTTTTTTT 3'  
Average insert size: 1.1 kb."  
/db\_xref="taxon:9606"  
/clone="IMAGE:1175743"  
/clone\_lib="NCI-CGAP\_Pr25"  
/tissue\_type="epithelium (cell line)"  
/lab\_host="SOLR (kanamycin resistant)"  
BASE COUNT 131 a 179 c 181 g 97 t 3 others  
ORIGIN  
Query Match 68.4%; Score 446; DB 11; Length 591;  
Best Local Similarity 95.4%; Pred. No. 0.00e+00;  
Matches 544; Conservative 0; Mismatches 11; Indels 15; Gaps 15;  
Db 28 AAATTCATGTTCCAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 87  
|||||  
Cp 652 AAATTCATGTTCCAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 593  
Db 88 GTGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCCCTGAC 147  
|||||  
Cp 592 GTGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCCCTGAC 533



Db 148 AGACTAGCAGGCCCGAGTGGTTCATGAAGCGTGACGGCCACCATCACATCACAAGCGTGATGGC 207  
|||||  
Cp 532 AGACTAGCAGGCCCGAGTGGTTCATGAAGCGTGACGGCCACCATCACATCACAAGCGTGATGGC 473  
Db 208 ACCTGCTGGTG-AGCTGCAGTCCCTGGTGGGCCGGGCGACACTCACTCTGGCAGCCGC-GTCC 265  
|||||  
Cp 472 ACCTGCTGGTGAGCTGCAGTCCCTGGTGGGCCGGGCGACACTCACTCTGGCAGCCCGCTCC 413  
Db 266 CCCACGCCAGCCATGTTGTCGCCGCATACGCAGTAGTTGAGTTGAGTGGTTCGCGAAGGCCAG 325  
|||||  
Cp 412 CCCACGCCAGCCATGTTGTCGCCGCATACGCAGTAGTTGAGTTGAGTGGTTCGTAAGGCCAG 353  
Db 326 CCAGGCCAGGTAGGGTAGAGCAGCGGGCGGCGGCGGCGGCTCACCTGGTACCAAGGCCAC 385  
|||||  
Cp 352 CCAGGCCAGGTAGGGTAGAGCAGCGGGCGGCGGCGGCGGCTCACCTGGTACCAAGGCCAC 293  
Db 386 GGTACTGGNCNC-GCCGCCCACTGACCAGCAGGAGATCCACCAAGGCCCGCCATTTG 444  
|||||  
Cp 292 GGTAGTGGTGCCTGCCCGCCCACTGACCAGCAGGAGATCCACCAAGGCCCGCCATTTG 233  
Db 445 TCGGGCAC-AAAGAAGATGGGGGCCATGCC-AGTTCAGGGCCAGCTGCC-AGTGTAGAN 501  
|||||  
Cp 232 TCGGGCACCAAGAGATGGGGGCCATGCCAGTTTACGGGCCAGCTGCCCGCTGTAGAG 173  
Db 502 GCC-AGGGGAAC-ACAGC-TTCTCTGTGAAGCCTCC-AGCTCTTTC-AGAC-AGGTAGGA 555  
|||||  
Cp 172 GCCAGGGGAACACAGCCCTTCTGTGAAGCCTCCAGCTCTTTCCAGACCAGGTAGGA 113  
Db 556 GC-GTACCC-ATG-CTGAGTAGAGCTTGCC 582  
||  
Cp 112 GCCGTACCCCATGGCTGAGTAGAGCGTGCC 83  
RESULT 8 AA210894 556 bp mRNA EST 13-AUG-1997  
LOCUS  
DEFINITION zr90g06.s1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:683002 3',  
similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR  
(HUMAN);, mRNA sequence.  
ACCESSION AA210894  
NID g1809540  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 556)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert\_Strausberg@nih.gov  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Seq primer: -41ml3 fwd. ET from Amersham  
High quality sequence stop: 428.  
Location/Qualifiers  
1. .556  
/organism="Homo sapiens"  
/note="vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was prepared from human tonsillar cells enriched for  
germinal center B cells by flow sorting (CD20+, IgD-),  
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was  
primed with a Not I - oligo(dT) primer  
[5'-TGTTACCAATCTGAAGTGGGAGCGCGCGCTCATTTTTTTTTTTTTTTT-  
3']. Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. Library

went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:683002"  
/clone\_lib="NCI\_CGAP\_GCB1"  
/tissue\_type="germinal center B cell"  
/lab\_host="DH10B"  
complement(<1..>556)  
/db\_xref="GDB:5586322"  
BASE COUNT 122 a 160 c 164 g 104 t 6 others  
ORIGIN  
mRNA  
Query Match 67.8%; Score 442; DB 8; Length 556;  
Best Local Similarity 96.9%; Pred. No. 0.00e+00;  
Matches 473; Conservative 0; Mismatches 11; Indels 4; Gaps 4;  
Db 43 AAATTCATGTTCCAAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCG 102  
|||||  
Cp 652 AAATTCATGTTCCAAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCG 593  
Db 103 GTGGGGCTCAGTGGGGCCACCTCTCTGAAGCTCTGCTGACCCCTGGGCCAAGCCNTGAC 162  
|||||  
Cp 592 GTGGGGCTCAGTGGGGCCACCTCTCTGAAGCTCTGCTGACCCCTGGGCCAAGCCCTGAC 533  
Db 163 AGACTAGCAGGCCCGAGTGGTTCATGAAGCGTACGGGCCACCATCATCACAAAGCGTGATGGC 222  
|||||  
Cp 532 AGACTAGCAGGCCCGAGTGGTTCATGAAGCGTACGGGCCACCATCATCACAAAGCGTGATGGC 473  
Db 223 ACCTGCTGTFN-AGCTGCAGTCCCTGGTGGGCCGGGCGACTCACTTCTGGCAGCCCGCTCC 281  
|||||  
Cp 472 ACCTGCTGTFGAGCTGCAGTCCCTGGTGGGCCGGGCGACTCACTTCTGGCAGCCCGCTCC 413  
Db 282 CCCACGCCAGCCATGTTGTCCCGCCATACGCAGTAGTTGAGTGTGGTTCGGAAGGCCAG 341  
|||||  
Cp 412 CCCACGCCAGCCATGTTGTCCCGCCATACGCAGTAGTTGAGTGTGGTTCGTAAGGCCAG 353  
Db 342 CCAGGCCAGGTAGGGGTAGAGCAGGGCGGGCGGCCAGCGGGCTCACTGGTACCAGGCCAC 401  
|||||  
Cp 352 CCAGGCCAGGTAGGGGTAGAGCAGGGCGGGCGGCCAGCGGGCTCACTGGTACCAGGCCAC 293  
Db 402 GGAGTGTGNCCTCCCGCCCGCCCTGACCAGCAGGAGATTCCACCAAGGCCCGCCATTTG 461  
|||||  
Cp 292 GGAGTGTGTCGCCCGCCCGCCCTGACCAGCAGGAGATCCACCAAGGCCCGCCATTTG 233  
Db 462 TCGGGCACCAAGAAGATGGGGGC-ATGCC-AGTTCAGGGCCAGCTGCC-AGTGTANAG 518  
|||||  
Cp 232 TCGGGCACCAAGAAGATGGGGGCCATGCCAGTTTACGGGCCAGCTGCCAGTGTAGAG 173  
Db 519 GCCAAGGG 526  
|||  
Cp 172 GCCCAGGG 165  
RESULT 9 AA479680 541 bp mRNA EST 09-NOV-1997  
LOCUS  
DEFINITION zu43a07.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 740724  
3', similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE  
RECEPTOR (HUMAN);, mRNA sequence.  
ACCESSION AA479680  
NID g2205566  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 541)  
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,  
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,  
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,  
Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.  
TITLE WashU-NCI human EST Project  
JOURNAL Unpublished (1997)  
COMMENT



Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estewatson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 598 Std Error: 0.00  
Seq primer: -41m13 fwd. ET from Amersham  
High quality sequence stop: 52.

FEATURES  
source  
1. .541  
/organism="Homo sapiens"  
/note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a  
modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer [5'  
TGTTACCAATCTGAAGTGGGAGCGCGGTTTTTTTTTTTTTTT 3'],  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified pT7T3 vector  
(Pharmacia). Library constructed by Bento Soares and  
M.Fatima Bonaldo."  
/db\_xref="GDB:5940628"  
/db\_xref="taxon:9606"  
/clone="740724"  
/clone\_lib="Soares ovary tumor NbHOT"  
/sex="Female"  
/tissue\_type="ovarian tumor"  
/lab\_host="DH10B (ampicillin resistant)"  
BASE COUNT 111 a 155 c 155 g 119 t 1 others  
ORIGIN

Query Match 67.0%; Score 437; DB 11; Length 541;  
Best Local Similarity 97.9%; Pred. No. 0.00e+00;  
Matches 461; Conservative 0; Mismatches 7; Indels 3; Gaps 3;  
Db 73 AAATTCCATGTTCCACAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 132  
|||||  
Cp 652 AAATTCCATGTTCCACAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 593  
Db 133 GTGGGGGCTCAGGTGGGCGCACCTCTGAAGCTCTGTGACCCCTGGGCCAAGGCCTCGAC 192  
|||||  
Cp 592 GTGGGGGCTCAGGTGGGCGCACCTCTGAAGCTCTGTGACCCCTGGGCCAAGGCCTGAC 533  
Db 193 AGACTAGCAGGCCCGCAGTGGTGCATGAAAGCGTGCAGCGCCACCATCATCACAAGCGTGATGC 252  
|||||  
Cp 532 AGACTAGCAGGCCCGCAGTGGTGCATGAAAGCGTGCAGCGCCACCATCATCACAAGCGTGATGC 473  
Db 253 ACCTGCTGGTGAAGCTCGAGTCCCTGGTGGGCGCGGCACCTCACTCTGGCAGCGCCGTCC 312  
|||||  
Cp 472 ACCTGCTGGTGCAGCTGCAGTCCCTGGTGGGCGCGGCACCTCACTCTGGCAGCGCCGTCC 413  
Db 313 CCCACGCCAGCCATGGTGTCCCGCCCATACGCAGTAGTTGAGTGTGGTCGCGAAGGCCAG 372  
|||||  
Cp 412 CCCACGCCAGCCATGGTGTCCCGCCCATACGCAGTAGTTGAGTGTGGTCGTAAGGCCAG 353  
Db 373 CCAGGCACGGTAGGGGTAGAGCAGGCGGGCGGCCAGCGGGCTCCACCTGGTACGAGGCCA 432  
|||||  
Cp 352 CCAGGCCAGGTAGGGGTAGAGCAGGCGGGCGGCCAGCGGGCTC-ACCTGGTACGAGGCCA 294  
Db 433 CGGTAGTGGCNGC-GCCGCGCCCACTGACCAGCAGGAGATCCACCAAGGCCCAAGCCCATTT 491  
|||||  
Cp 293 CGGTAGTGGTGCCTGCGCGCGCCCACTGACCAGCAGGAGATCCACCAAGGCCCAAGCCCATTT 234  
Db 492 GTCGGGCACCAAGAAGATGGGGGCCCATGCC-AGTTCAGGGGCCAGCTGCC 541  
|||||  
Cp 233 GTCGGGCACCAAGAAGATGGGGGCCCATGCCAGTTCAGGGGCCAGCTGCC 183

RESULT 10  
LOCUS AA748434 541 bp mRNA EST 18-FEB-1998  
DEFINITION ny01b10.s1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:1270459

similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR  
(HUMAN);, mRNA sequence.  
ACCESSION AA748434  
NID 92788392  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 541)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert\_Strausberg@nih.gov  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,  
Ph.D., Gerald Marti, M.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 874 Std Error: 0.00  
Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 358.  
FEATURES  
Location/Qualifiers  
1. .541  
/organism="Homo sapiens"  
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was prepared from human tonsillar cells enriched for  
germinal center B cells by flow sorting (CD20+, IgD-),  
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was  
primed with a Not I - oligo(dT) primer  
[5'-TGTTACCAATCTGAAGTGGGAGCGCGCCCTCATTTTTTTTTTTTTT-  
3']. Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT7T3 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."  
/db\_xref="taxon:9606"  
/clone="IMAGE:1270459"  
/clone\_lib="NCI\_CGAP\_GCB1"  
/tissue\_type="germinal center B cell"  
/lab\_host="DH10B"

BASE COUNT 120 a 152 c 166 g 100 t 3 others  
ORIGIN  
Query Match 65.2%; Score 425; DB 15; Length 541;  
Best Local Similarity 94.7%; Pred. No. 0.00e+00;  
Matches 463; Conservative 0; Mismatches 23; Indels 3; Gaps 3;  
Db 50 AAATTCCATGTTCCAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 109  
|||||  
Cp 652 AAATTCCATGTTCCAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 593  
Db 110 GTGGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCCCTGAC 169  
|||||  
Cp 592 GTGGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCCCTGAC 533  
Db 170 AGACTAGCAGGCCCGCAGTGGTGCATGAAAGCGTGACGGCCACCATCATCACAAGCGTGATGC 229  
|||||  
Cp 532 AGACTAGCAGGCCCGCAGTGGTGCATGAAAGCGTGACGGCCACCATCATCACAAGCGTGATGC 473  
Db 230 ACCTGCTGGTGCAGCTGCAGTCCCTGGTGGGCGCGGCACCTCACTCTGGCAGCCGC-GTCC 288  
|||||

Cp 472 ACCTGCTGGTGCAGCTGCAGTCCCTGGTGGCGCGGCACCTCACTCTGGCAGCCCGCGTCC 413

Db 289 CCCAGCCAGCCATGGTTGTCCCGCCATACGCAGTAGTTGAGTGTGGTCGCGAAGGCCAG 348  
|||||

Cp 412 CCCAGCCAGCCATGGTTGTCCCGCCATACGCAGTAGTTGAGTGTGGTCGTAAGGCCAG 353  
|||||

Db 349 CCAGGCCAGGTAGGGGTAGAGCAGGGCGGCGCCAGCGGGCTGCACCTGGTAGCAGGCCA 408  
|||||

Cp 352 CCAGGCCAGGTAGGGGTAGAGCAGGGCGGCGCCAGCGGGCT-CACCTGGTACCAGGCCA 294  
|||||

Db 409 CGGTAGTGNCCGCGCCGCCACTGACCAGCAGGAGATTACCAAGGCCCCAGCCATT 468  
|||||

Cp 293 CGGTAGTGGTGCAGCGCGCCGCCACTGACCAGCAGGAGATCCACCAAGGCCCCAGCCATT 234  
|||||

Db 469 GTTCGGNAGCAAGAAGATGGGGGGCATGAC-ATTTAAGGGCAAGTTGCCCAGTGAGA 527  
|| |||

Cp 233 GTCGGGCACCAAGAAGATGGGGGGCATGCCAGTTCAGGGCCAGCTGCCCAGTGAGA 174  
|| |||

Db 528 GGGCCAGGG 536  
|| |||||

Cp 173 GGCCCAGGG 165

RESULT 11

LOCUS AA456315 506 bp mRNA EST 06-JUN-1997

DEFINITION aal3g10.s1 Soares NhMPu S1 Homo sapiens cDNA clone 813186 3', similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR (HUMAN);, mRNA sequence.

ACCESSION AA456315

NID 92179525

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 506)

AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Gelsel,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.

TITLE WashU-Merck EST Project 1997

JOURNAL Unpublished (1997)

COMMENT

Contact: Wilson RK

WashU-Merck EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -41m13 fwd. ET from Amersham

High quality sequence stop: 443.

Location/Qualifiers

1..506

/organism="Homo sapiens"

/note="Organ: mixed (see below); Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbHM, pregnant uterus NbHPU, and fetal heart NbHH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

/db\_xref="taxon:9606"

/clone="813186"

/clone\_lib="Soares NhMPu S1"

/tissue\_type="Pooled human melanocyte, fetal heart, and

pregnant uterus" /lab\_host="DH10B" complement(<1..>506) /db\_xref="GDB:6044096"

BASE COUNT 114 a 150 c 158 g 83 t 1 others

ORIGIN

Query Match 65.0%; Score 424; DB 7; Length 506; Best Local Similarity 96.9%; Pred. No. 0.00e+00; Matches 463; Conservative 0; Mismatches 10; Indels 5; Gaps 5;

Db 34 AAATTCCATGTTCCAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACACTGCTCCCG 93  
|||||

Cp 652 AAATTCCATGTTCCAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACACTGCTCCCG 593  
|||||

Db 94 GTGGGGCTCAGGTGGGGCAACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGCCCTGAC 153  
|||||

Cp 592 GTGGGGCTCAGGTGGGGCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGCCCTGAC 533  
|||||

Db 154 AGACTAGCAGGCCCATGGTTCATGAAGCGTGACGGCCACCATCATCACAAGCGTGATGGC 213  
|||||

Cp 532 AGACTAGCAGGCCCATGGTTCATGAAGCGTGACGGCCACCATCATCACAAGCGTGATGGC 473  
|||||

Db 214 ACCTGCTGGTGAGGCTGCAGTCCCTGGTGGCGCGGCACCTCACTCTGGCAGCCCGTCC 273  
|||||

Cp 472 ACCTGCTGGTGCAGCTGCAGTCCCTGGTGGCGCGGCACCTCACTCTGGCAGCCCGTCC 413  
|||||

Db 274 CCCATGCCAGCCATGGTTGTCCCGCATACGCAGTAGTTAGTGTGGTCGCGAAGGCCAG 333  
|||||

Cp 412 CCCAGCCAGCCATGGTTGTCCCGCATACGCAGTAGTTAGTGTGGTCGTAAGGCCAG 353  
|||||

Db 334 CCAGGCCAGGTAGGGGTAGAGCAGCGGGCGGCCAGCGGGCTACCTGGTACCAGGCCAC 393  
|||||

Cp 352 CCAGGCCAGGTAGGGGTAGAGCAGCGGGCGGCCAGCGGGCTACCTGGTACCAGGCCAC 293  
|||||

Db 394 GGTAAGTGGG-CCGCGCGCCACTGACCAGCAGGAGATCCACCAAGGCCAGCC-ATTG 451  
|||||

Cp 292 GGTAAGTGGTCCGCGCGCCACTGACCAGCAGGAGATCCACCAAGGCCAGCCATTG 233  
|||||

Db 452 TCGGGCAC-AAAGAAGATGGGGGCGCATGCC-AGTTTCAGAGGCGAGCTGCC-AGTGTAG 506  
|||||

Cp 232 TCGGGCACAAAGAAGATGGGGGCGCATGCCAGTTCAGSGCCAGCTGCCCAGTGTAG 175  
|||||

RESULT 12

LOCUS AA036726 551 bp mRNA EST 10-MAY-1997

DEFINITION zK30b07.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 472021 3' similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR (HUMAN);, mRNA sequence.

ACCESSION AA036726

NID g1509983

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 551)

AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.

TITLE WashU-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT

Contact: Wilson RK

WashU-Merck EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 695 Std Error: 0.00  
Seq primer: -40M13 fwd. from Amersham  
High quality sequence stop: 321.

FEATURES  
source

1. .551  
/organism="Homo sapiens"  
/note="Organ: uterus; Vector: pT7T3-Pac; Site\_1: Not I;  
Site\_2: Eco RI; 1st strand cDNA was primed with a Not I -  
oligo(dT) primer [5',  
AACTGGAAGAATTCCGGCGCCCTTTTTTTTTTTTTTTT 3'],  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT7T3 vector. Library  
went through one round of normalization. Library  
constructed by M. Fatima Bonaldo."  
/db\_xref="taxon:9606"  
/clone="472021"  
/clone\_lib="Soares pregnant uterus NbHPU"  
/sex="female"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
complement(<1.>551)  
/db\_xref="GDB:3758023"

mrna

BASE COUNT 123 a 160 c 164 g 96 t 8 others

ORIGIN

Query Match 63.5%; Score 414; DB 5; Length 551;  
Best Local Similarity 93.5%; Pred. No. 0.00e+00;  
Matches 491; Conservative 0; Mismatches 24; Indels 10; Gaps 10;

Db 30 AAATTCCATGTTCCAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 89  
|||||  
Cp 652 AAATTCCATGTTCCAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 593  
|||||  
Db 90 GTGGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCTG-GAC 148  
|||||  
Cp 592 GTGGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCTGAC 533  
|||||  
Db 149 AGACTAGCAGGCCAGTGGTCATGAAAGCGTGCAGCGGCCACCACATCATCAAGCGTGATGGC 208  
|||||  
Cp 532 AGACTAGCAGGCCAGTGGTCATGAAAGCGTGCAGCGGCCACCACATCATCAAGCGTGATGGC 473  
|||||  
Db 209 ACCTGCTGGT-CAGCTGCAGTCCCTGGTGGGGGGGCACTCACTCTGGCAGCCGC-GTCC 266  
|||||  
Cp 472 ACCTGCTGGTGCAGTGCAGTCCCTGGTGGGGGGGCACTCACTCTGGCAGCCGCTCC 413  
|||||  
Db 267 CCCACGCCAGCCATGGTTGTCCCGCCATACGCAGTAGTTGAGTGTGGTGGTGAAGGCCAG 326  
|||||  
Cp 412 CCCACGCCAGCCATGGTTGTCCCGCCATACGCAGTAGTTGAGTGTGGTGGTGAAGGCCAG 353  
|||||  
Db 327 CCAGGCANGGTAAGGGGTAGAGCAGGCGNGGGGCGCAGCGGGCTCACCTGGTACCAGGCCA 386  
|||||  
Cp 352 CCAGGCCAGGTA-GGGGTAGAGCAGGCGGGGCGGCGCAGCGGGCTCACCTGGTACCAGGCCA 294  
|||||  
Db 387 CGGTACTGTGANCAGCCGCNCCACTTGACCAGCAGGAGATCCACCAGGCCAGNCCATT 446  
|||||  
Cp 293 CGGTAGTGGTGCCTGCCCGCCCCCACT-GACCAGCAGGAGATCCACCAGGCCAGCCCAT 235  
|||||  
Db 447 TTGTTCCGGGCACCAAGAAGATGGGGGGCCATGCCAGTTTCAGGGGCCAAGCTGCCATGTN 506  
|||||  
Cp 234 T-GT-CGGGCACCAAGAAGATGGGGGGCCATGCCAGTTTCAGGGGCCA-GCTGCCCACT- 179  
|||||  
Db 507 GTAGAGGCCCCAGGGGAACACAGNCTTCTTTGTGANGCCTTCCA 551  
|||||  
Cp 178 GTAGAGGCCCCAGGGGAACACAGCCTTCTCTGT-GAAGCCTTCCCA 135  
|||||

RESULT 13  
LOCUS AA838533 505 bp mrna EST 18-MAR-1998  
DEFINITION oe39a11.s1 NCI\_CGAP\_Pr25 Homo sapiens cDNA clone IMAGE:1410908  
similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR  
(HUMAN);, mRNA sequence.

ACCESSION

AA838533

NID

g2913332

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 505)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert\_Strausberg@nih.gov

Tissue Procurement: Suzanne L. Topalian, M.D., Robert K. Bright,

Ph.D.

cDNA Library Preparation: Stratagene, Inc.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 894 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 323.

FEATURES

Location/Qualifiers

source

1. .505

/organism="Homo sapiens"

/note="Organ: prostate; Vector: Bluescript SK-; Site\_1:

EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer:

Oligo dT. Normal prostate epithelial cell line (HPV

immortalized). 5' adaptor sequence: 5' GAATTCGGCAGCAG 3'

3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3'

Average insert size: 1.1 kb."

/db\_xref="taxon:9606"

/clone="IMAGE:1410908"

/clone\_lib="NCI\_CGAP\_Pr25"

/tissue\_type="epithelium (cell line)"

/lab\_host="SOLR (kanamycin resistant)"

112 a 148 c 156 g 89 t

BASE COUNT

ORIGIN

Query Match 62.1%; Score 405; DB 16; Length 505;

Best Local Similarity 94.9%; Pred. No. 0.00e+00;

Matches 444; Conservative 0; Mismatches 21; Indels 3; Gaps 3;

Db 41 AAATTCCATGTTCCAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 100

|||||

Cp 652 AAATTCCATGTTCCAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 593

|||||

Db 101 GTGGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCTTGAC 160

|||||

Cp 592 GTGGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCTTGAC 533

|||||

Db 161 AGACTAGCAGGCCAGTGGTTCATGAAGCGTGACGGCCACCACATCATCAAGCGTGATGGC 220

|||||

Cp 532 AGACTAGCAGGCCAGTGGTTCATGAAGCGTGACGGCCACCACATCATCAAGCGTGATGGC 473

|||||

Db 221 ACCTGCTGGTCAGCTGCAGTCCCTGGTGGGGCCGGGCACTCACTCTGGCAGCCCGCTCC 280

|||||

Cp 472 ACCTGCTGGTCAGCTGCAGTCCCTGGTGGGGCCGGGCACTCACTCTGGCAGCCCGCTCC 413

|||||

Db 281 CCCACGCCAGCCATGGTTGTCCCGCCATACGCAGTAGTTGAGTGTGGTGCGAAG-CCAG 339

|||||

Cp 412 CCCACGCCAGCCATGGTTGTCCCGCCATACGCAGTAGTTGAGTGTGGTGCGAAGCCAG 353

|||||

Db 340 CCAGGCAAGGTAGGGGTAGAGCAGGCGGGCGGGC-AGCGGGCTCACTGGTACCAGGCCAC 398

|||||

Cp 352 CCAGGCAAGGTAGGGGTAGAGCAGGCGGGCGGGCAGCGGGCTCACTGGTACCAGGCCAC 293

|||||







```
/clone="IMAGE:1324464"  
/clone_lib="NCI_CGAP_Kid5"  
/tissue_type="2 pooled tumors (clear cell type)"  
/lab_host="DH10B"
```

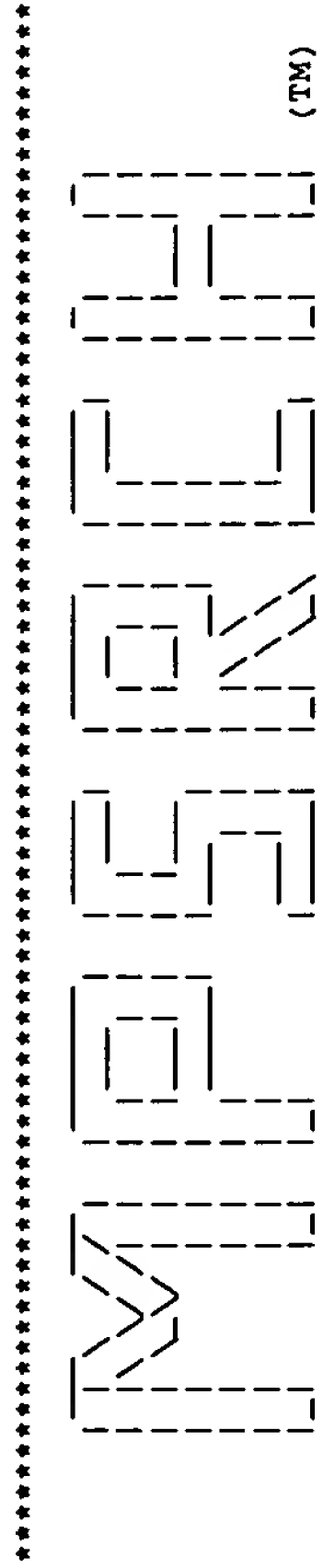
```
BASE COUNT      111 a      144 c      151 g      101 t  
ORIGIN
```

```
Query Match          60.4%; Score 394; DB 13; Length 507;  
Best Local Similarity 96.5%; Pred. No. 0.00e+00;  
Matches 435; Conservative 0; Mismatches 11; Indels 5; Gaps 5;
```

```
Db 53 AAATTCCATGTTCCAAAGACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 112  
|||||  
Cp 652 AAATTCCATGTTCCAAAGACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 593  
|||||  
Db 113 GTGGGGCTCAGGTGGGCCACCTCTGAAGCTCTGTGACCCCTGGGCCAAGGCTGTAC 172  
|||||  
Cp 592 GTGGGGCTCAGGTGGGCCACCTCTGAAGCTCTGTGACCCCTGGGCCAAGGCTGTAC 533  
|||||  
Db 173 AGACTAGCAGGCCCCAGTGGTCATGAAAGCGTGACGGCCACCACATCACAAAGCGTGATGGC 232  
|||||  
Cp 532 AGACTAGCAGGCCCCAGTGGTCATGAAAGCGTGACGGCCACCACATCACAAAGCGTGATGGC 473  
|||||  
Db 233 ACCTGCTGGTGAGTCTGAGTCCCTGGTGGCCGGGCACTCACTCTGGCAGCGCCGTCC 292  
|||||  
Cp 472 ACCTGCTGGTGAGTCTGAGTCCCTGGTGGCCGGGCACTCACTCTGGCAGCGCCGTCC 413  
|||||  
Db 293 CCCACGCCAGCCATGTTGTCCCGCCATACGCAGTAGTTGAGTGTGGTGTGAAGGCCAG 352  
|||||  
Cp 412 CCCACGCCAGCCATGTTGTCCCGCCATACGCAGTAGTTGAGTGTGGTGTGAAGGCCAG 353  
|||||  
Db 353 CCAGGCCAGGTAGGGGTAGAGCAGGGCGGGCCAGCGGGCTACACCTGGTACCAGGCCA 412  
|||||  
Cp 352 CCAGGCCAGGTAGGGGTAGAGCAGGGCGGGCCAGCGGGCT-CACCTGGTACCAGGCCA 294  
|||||  
Db 413 CGGTAGGTGACAC-GCCGCCCCACTGACCAGCAGGAGATCCAC-AAGGCC-AGCCCCATTT 469  
|||||  
Cp 293 CGGTAGTGGCTGCCGCCGCCCTGACCAAGCAGGAGATCCACCAAGGCCCCAGCCCCATTT 234  
|||||  
Db 470 GTCGGGCACCAA-GAAGATGGGGGCCCATGC 499  
|||||  
Cp 233 GTCGGGCACCAAAGAGATGGGGGCCCATGC 203  
|||||
```

Search completed: Tue Jul 21 15:24:46 1998  
Job time : 776 secs.

**THIS PAGE BLANK (USPTO)**



\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mpsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Jul 21 16:05:14 1998; MasPar time 108.16 Seconds  
767.738 Million cell updates/sec

Tabular output not generated.

Title: >US-09-047-652A-2  
Description: (1-652) from US09047652A.seq  
Perfect Score: 652  
N.A. Sequence: 1 CCACGGCGAGGGTCTCCGCT.....GTTCTTGAACATGGAATTT 652  
Comp: GGTGCCGCTCCAGAGGCGA.....CAAGAACCTGTACCTTAA

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 176886 seqs, 63680241 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-geneseq31-2

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37

Statistics: Mean 8.516; Variance 5.456; scale 1.561

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	238	36.5	272	20	T22195	Human gene signature	6.34e-134
c 2	44	6.7	91	9	Q51746	Oligonucleotide probe	2.41e-11
c 3	41	6.3	204	1	N81164	Base substituted E.co	1.11e-09
4	40	6.1	91	9	Q51746	Oligonucleotide probe	3.92e-09
c 5	36	5.5	114	12	Q70465	Generic DNA sequence	5.68e-07
6	36	5.5	204	1	N81164	Base substituted E.co	5.68e-07
7	35	5.4	114	12	Q70469	Generic DNA sequence	1.93e-06
8	34	5.2	114	12	Q70468	Generic DNA sequence	6.48e-06
9	34	5.2	114	12	Q70472	Generic DNA sequence	6.48e-06
10	34	5.2	114	12	Q70465	Generic DNA sequence	6.48e-06
11	34	5.2	114	12	Q70467	Generic DNA sequence	6.48e-06
c 12	34	5.2	114	12	Q70468	Generic DNA sequence	6.48e-06
c 13	34	5.2	114	12	Q70467	Generic DNA sequence	6.48e-06
c 14	33	5.1	114	12	Q70469	Generic DNA sequence	2.16e-05





```
RESULT 5
ID Q70465 standard; DNA; 114 BP.
AC Q70465;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR petide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PSDB; R65150 and R65151.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PT Disclosure; Page 35; 255pp; English.
CC Q70465 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)6(TGC)(NNB)11Z(NNB)14(TGC)(NNB)3Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or comsps. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed
CC activity allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 5.5%; Score 36; DB 12; Length 114;
Best Local Similarity 4.5%; Pred. No. 5.68e-07;
Matches 5; Conservative 33; Mismatches 74; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbntgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnn 62
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 151 CTCTGTGAAGCCTCCAGCTCTTCCAGACCAGGTAGGAGCCGTACCCCATGGCTGAGTA 92
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbntgcnnbnnbnnb 114
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 91 GAGCGTGCCCCAGACAGGGCCCCAGCACCCAGTCGGGGCGGTGCCACGAGGGC 40

RESULT 6
ID N81164 standard; DNA; 204 BP.
AC N81164;
DT 08-NOV-1990 (first entry)
DE Base substituted E.coli beta-galactosidase alpha-fragment.
KW E.coli beta galactosidase alpha-fragment; base substitutions; ss.
```

```
OS Escherichia coli.
FH Key Location/Qualifiers
FT misc_feature 19..69
FT /*tag= a
FT /function= multiple cloning site
FT primer_bind 187..204
FT /*tag= b
PN EP-285123-A.
PD 05-MAY-1988.
PF 30-MAR-1988; 105163.
PR 03-APR-1987; US-034819.
PA (SUSO) SUOMEN SOKERI OY.
PI Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
DR WPI; 88-279927/40.
PT Introducing random point mutations into nucleic acids -
PT by prepn of single stranded template, annealing a primer, elongation,
PT misincorporation, completion of molecules and screening.
PS Disclosure; p; English.
CC Random point mutations were introduced into the alpha fragment of
CC E.coli beta-galactosidase. The wild type sequence was obtained as a
CC single stranded template and an oligonucleotide was hybridised to
CC it to generate a popn of DNA molecules which terminate at all
CC possible nucleotide positions within a specified region. The
CC variable 3' ends generated in this way are used as primers for
CC reverse transcriptase. Nucleotides are misincorporated by the
CC transcriptase and the molecules are completed to forms that can be
CC amplified and then expressed in a suitable host-vector system.
CC The sequence covers all 176 diff base substitutions, most of which
CC occurred singularly in any given mutant.
SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 5.5%; Score 36; DB 1; Length 204;
Best Local Similarity 15.1%; Pred. No. 5.68e-07;
Matches 16; Conservative 53; Mismatches 35; Indels 2; Gaps 2;

Db 93 hyrrmrbnvyrdyn-rsdaaawccyrrsvkydcccynachhddhyv-ybbbvynvnhnn 150
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 19 CTGGTACGGCGCCCTGCAGAGCCCTCGTGGCACCCGCCACTGGTGCTGGCCCTGT 78
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 151 cnccebnnhvhvbnhnnhrnwayvrhdarrddvhccvchccgac 196
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 79 CTGGGGCAGGCTCTACTCAGCCATGGGGTACGGCTCCTACCTGGTC 124

RESULT 7
ID Q70469 standard; DNA; 114 BP.
AC Q70469;
DT 07-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6,9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PT Disclosure; Page 35; 255pp; English.
```

CC Q70469 is a generic DNA sequence used to generate random TSAR peptide  
CC This generic formula can be represented as follows: X(TGC)(NNB)10-  
CC (TGC)(NNB)6Z(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction  
CC sites (X is not the same as Y) that are not specified further. Other  
CC sequence generates peptides that are cloverleaf in structure. Other  
CC generic sequences are shown in Q70465-68. Other specific peptides  
CC generated by these generic sequences are shown in R65150-54. TSARs are  
CC concatenated heterofunctional proteins or peptides, comprising at least  
CC two functional regions - a binding domain with affinity for a ligand and  
CC a second effector peptide portion that is chemically or biologically  
CC active.They may further comprise a linker peptide between the 2 domains.  
CC The oligonucleotides are also designed so that the expressed peptide  
CC contains 2 or 4 cysteine residues positioned in, or flanking, the  
CC unpredicted or variant residues. These residues confer some degree of  
CC conformational rigidity to the peptides. The TSARs or compsns. comprising  
CC a TSAR binding domain can be used in vivo to deliver a chemically or  
CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin  
CC or enzyme, to the specific target or on the cell. They can also replace  
CC the function of macromolecules, eg. monoclonal or polyclonal antibodies  
CC and therefore circumvent the need for complex methods of hybridoma  
CC formation or in vivo antibody production. The TSARs are easily  
CC characterised and have designed activity allowing direct and rapid  
CC detection in a screening process.  
SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;  
  
Query Match 5.4%; Score 35; DB 12; Length 114;  
Best Local Similarity 6.3%; Pred. No. 1.93e-06;  
Matches 7; Conservative 31; Mismatches 73; Indels 0; Gaps 0;  
  
Db 3 cnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnn 62  
QY 16 CCGCTGGTACGCCGGCCTGCAGAGCCCTCGTGGCAGCCGCCGCCACTGGGTGCTGGGCC 75  
Db 63 bnnbtgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbtg 113  
QY 76 TGTCTGGGCACGCTCTACTCAGCCATGGGGTACGGCTCCTACCTGGTCTG 126  
  
RESULT 8  
ID Q70468 standard; DNA; 114 BP.  
AC Q70468;  
DT 05-APR-1995 (first entry)  
DE Generic DNA sequence to generate a random TSAR petide library.  
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;  
KW effector domain; concatenated heterofunctional protein; linker;  
KW direct; rapid; detection; screening; treatment; generic; ss.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT misc\_feature 55..60  
FT /\*tag= a  
FT /note= "this sequence represents 'Z'; Z can be a  
FT sequence of 6, 9 or 12 nucleotides (see  
FT comments)"  
PN WO9418318-A.  
PD 18-AUG-1994.  
PF 01-FEB-1994; U00977.  
PR 01-FEB-1993; US-013416.  
PR 30-DEC-1993; US-176500.  
PR 31-JAN-1994; US-189331.  
PA (UYNC-) UNIV NORTH CAROLINA.  
PI Fowlkes DM, Kay BK;  
DR WPI; 94-279739/34.  
DR P-PSDB; R65154.  
PT Identifying proteins or peptide(s) which bind a ligand - by  
PT screening a recombinant vector library expressing fusion proteins  
PT comprising a binding domain and an effector domain  
PS Disclosure; Page 35; 255pp; English.  
CC Q70468 is a generic DNA sequence used to generate random TSAR (Totally  
CC Synthetic Affinity Reagents) peptides.This generic formula can also be  
CC represented as follows: X(NNB)11(TGC)(NNB)6Z(NNB)7(TGC)(NNB)10Y. X  
CC and Y are flanking restriction sites (X is not the same as Y) that are  
CC not specified further. Other generic sequences are shown in Q70466-68.  
CC Other specific peptides generated by these generic sequences are shown in

CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,  
CC comprising at least two functional regions - a binding domain with  
CC affinity for a ligand and a second effector peptide portion that is  
CC chemically or biologically active.They may further comprise a linker  
CC peptide between the 2 domains.The oligonucleotides are also designed so  
CC that the expressed peptide contains 2 or 4 cysteine residues positioned  
CC in, or flanking, the unpredicted or variant residues. These residues  
CC confer some degree of conformational rigidity to the peptides. The TSARs  
CC or compsns. comprising a TSAR binding domain can be used in vivo to  
CC deliver a chemically or biologically active moiety, eg. metal ion,  
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the  
CC cell. They can also replace the function of macromolecules, eg.  
CC monoclonal or polyclonal antibodies and therefore circumvent the need  
CC for complex methods of hybridoma formation or in vivo antibody  
CC production. The TSARs are easily characterised and have designed activity  
CC allowing direct and rapid detection in a screening process.  
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;  
  
Query Match 5.2%; Score 34; DB 12; Length 114;  
Best Local Similarity 2.7%; Pred. No. 6.48e-06;  
Matches 3; Conservative 34; Mismatches 75; Indels 0; Gaps 0;  
  
Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnnnnnnn 62  
QY 16 CCGCTGGTACGCCGGCCTGCAGAGCCCTCGTGGCAGCCGCCGCCACTGGGTGCTGGGCC 75  
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnnbnnbnnb 114  
QY 76 TGTCTGGGCACGCTCTACTCAGCCATGGGGTACGGCTCCTACCTGGTCTGG 127  
  
RESULT 9  
ID Q70472 standard; DNA; 114 BP.  
AC Q70472;  
DT 10-APR-1995 (first entry)  
DE Generic DNA sequence to generate a random TSAR peptide library.  
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;  
KW effector domain; concatenated heterofunctional protein; linker;  
KW direct; rapid; detection; screening; treatment; generic; ss.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT misc\_feature 55..60  
FT /\*tag= a  
FT /note= "encoded by Z"  
PN WO9418318-A.  
PD 18-AUG-1994.  
PF 01-FEB-1994; U00977.  
PR 01-FEB-1993; US-013416.  
PR 30-DEC-1993; US-176500.  
PR 31-JAN-1994; US-189331.  
PA (UYNC-) UNIV NORTH CAROLINA.  
PI Fowlkes DM, Kay BK;  
DR WPI; 94-279739/34.  
DR P-PSDB; R58383.  
PT Identifying proteins or peptide(s) which bind a ligand - by  
PT screening a recombinant vector library expressing fusion proteins  
PT comprising a binding domain and an effector domain  
PS Disclosure; Page 36; 255pp; English.  
CC Q70472 is a generic DNA sequence used to generate random TSAR (Totally  
CC Synthetic Affinity Reagents) peptides.This generic formula can also be  
CC represented as follows: X(NNB)11(CAC)(NNB)11(CAC)(NNB)2Z(NNB)6  
CC -(CAC)(NNB)5(CAC)2(NNB)4. X and Y are flanking restriction sites  
CC (X is not the same as Y) that are not specified further. The peptides  
CC generated by this and other generic sequences (Q70470-73) have invariant  
CC histidine residues incorporated into variant sequences. TSARs are  
CC concatenated heterofunctional proteins or peptides, comprising at least  
CC two functional regions - a binding domain with affinity for a ligand and  
CC a second effector peptide portion that is chemically or biologically  
CC active.They may further comprise a linker peptide between the 2 domains.  
CC The TSARs or compsns. comprising a TSAR binding domain can be used in  
CC vivo to deliver a chemically or biologically active moiety, eg. metal  
CC ion, radioisotope, peptide, toxin or enzyme, to the specific target or  
CC on the cell. They can also replace the function of macromolecules, eg.









FH	Key	Location/Qualifiers
FT	misc_feature	55..60
FT		/*tag= a
FT		/note= "this sequence represents 'z'; 2 can be a
FT		sequence of 6, 9 or 12 nucleotides (see
FT		comments) *

PN	WO9418318-A.	
PD	18-AUG-1994.	
PF	01-FEB-1994;	U00977.
PR	01-FEB-1993;	US-013416.
PR	30-DEC-1993;	US-176500.
PR	31-JAN-1994;	US-189331.
PA	(UYNC-) UNIV	NORTH CAROLINA.

PR 01-FEB-1993; US-013416.  
PR 30-DEC-1993; US-176500.  
PR 31-JAN-1994; US-189331.  
PA (UYNC-) UNIV NORTH CAROLINA.  
PI Fowlkes DM, Kay BK;  
DR WPI; 94-279739/34.  
PT Identifying proteins or peptide(s) which bind a ligand - by  
PT screening a recombinant vector library expressing fusion proteins  
PT comprising a binding domain and an effector domain  
PS Disclosure; Page 35; 255pp; English.

PT Identifying proteins or peptide(s) which bind a ligand

Query Match 5.1%; Score 33; DB 12; Length 114;  
Best Local Similarity 5.6%; Pred. NO. 2.16e-05;  
Matches 6; Conservative 30; Mismatches 71; Indels 0; Gaps 0;

[illegible]

RESULT 15  
ID Q70466 standard; DNA; 114 BP.  
AC Q70466;  
DT 05-APR-1995 (first entry)  
DE Generic DNA sequence to generate a random TSAR-9 peptide library.  
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;  
KW effector domain; concatenated heterofunctional protein; linker;  
KW direct; rapid; detection; screening; treatment; generic; ss.  
OS Synthetic.

FT	misc_feature	Location/Qualifiers
FT	Key	55..60
FT		/*tag= a
FT		/note= "this sequence represents 'z'; z can be a
FT		sequence of 6, 9 or 12 nucleotides (see
FT		comments)"

PN WO9418318-A.

18-AUG-1994.  
01-FEB-1994; U00977.  
01-FEB-1993; US-013416.  
30-DEC-1993; US-176500.  
31-JAN-1994; US-189331.  
(UYNC-) UNIV NORTH CAROLINA.  
Fowlkes DM, Kay BK;  
WPI; 94-279739/34.  
P-PSDB; R65152.  
Identifying proteins or peptide(s) which bind a ligand - by  
screening a recombinant vector library expressing fusion proteins  
comprising a binding domain and an effector domain  
Disclosure; Page 35; 255pp; English.  
Q70466 is a generic DNA sequence used to generate random TSAR (Totally  
Synthetic Affinity Reagents) peptides. This generic formula can also be  
represented as follows: X(NNB)1(TGC)2(NNB)4Z(NNB)8(TGC)(NNB)  
-9Y. X and Y are flanking restriction sites (X is not the same as Y)  
that are not specified further. Other generic sequences are shown in  
Q70466-68. Other specific peptides generated by these generic sequences  
are shown in R65151-54. TSARS are concatenated heterofunctional proteins  
or peptides, comprising at least two functional regions - a binding  
domain with affinity for a ligand and a second effector peptide portion  
that is chemically or biologically active. They may further comprise a  
linker peptide between the 2 domains. The oligonucleotides are also  
designed so that the expressed peptide contains 2 or 4 cysteine residues  
positioned in, or flanking, the unpredicted or variant residues. These  
residues confer some degree of conformational rigidity to the peptides.  
The TSARS or compsns. comprising a TSAR binding domain can be used in  
vivo to deliver a chemically or biologically active moiety, eg. metal  
ion, radioisotope, peptide, toxin or enzyme, to the specific target or  
on the cell. They can also replace the function of macromolecules, eg.  
monoclonal or polyclonal antibodies and therefore circumvent the need for  
complex methods of hybridoma formation or in vivo antibody production.  
The TSARS are easily characterised and have designed activity allowing  
direct and rapid detection in a screening process.  
Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

Sequence 114 BP: 0 A: 4 C: 4 G: 4 T: 4

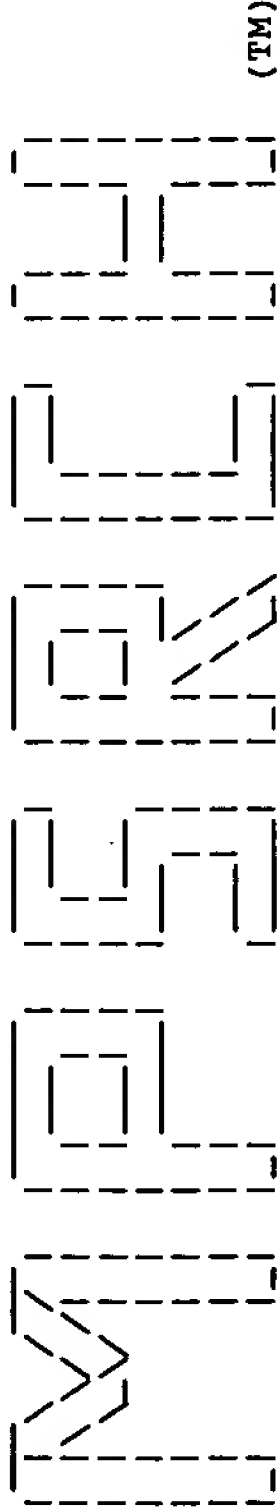
Query Match 4.9%; Score 32; DB 12; Length 114;  
Best Local Similarity 8.2%; Pred. NO. 7.10e-05;  
Matches 9; Conservative 28; Mismatches 73; Indels 0; Gaps 0;

Db	5	gcnnbnbnbnbnbnbnbnbnbnbnbnbnbtgctgcnnbnbnbnbnbnbnbnbnbnbnbnbn	64
		: : : : : : : :        : : :	
QY	191	GCCCTGAACGGGCATGCCCCCCCATCTTCTTTGGTGCCCGACAATAATGGCTGGGCGCTG	250
Db	65	nbnnbnbnbnbnbnbnbnbtgcnnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbn	114
		: : : : :     : : : : : : :	
QY	251	GTGGATCTCCTGCTGGTCAGTGGGGCGCGGCAGCCACFACCCTGGCCTG	300

Search completed: Tue Jul 21 16:07:07 1998  
Job time : 113 secs.

**THIS PAGE BLANK (USPTO)**

\*\*\*\*\*



\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Jul 21 16:07:26 1998; MasPar time 36.85 Seconds  
Tabular output not generated. 949.152 Million cell updates/sec

Title: >US-09-047-652A-2  
Description: (1-652) from US09047652A.seq  
Perfect Score: 652  
N.A. Sequence: 1 CCACGGCGAGGGTCTCCGCT.....GTTCTTGGAAACATGGAATTT 652  
Comp: GGTCGGCTCCAGAGCGGA.....CAAGAACCTGTACCTTAA

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 104157 seqs, 26825796 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-issued  
1:5\_COMB 2:PCT9\_COMB 3:backfiles

Statistics: Mean 8.110; Variance 4.794; scale 1.692

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	39	6.0	7218	1	US-08-232- Sequence 14, Applicati	2.64e-10
2	36	5.5	215	1	US-08-238- Sequence 5, Applicatio	1.55e-08
3	35	5.4	215	1	US-08-238- Sequence 5, Applicatio	5.93e-08
4	25	3.8	74	2	PCT-US95-1 Sequence 92, Applicati	2.03e-02
5	25	3.8	81	2	PCT-US95-1 Sequence 94, Applicati	2.03e-02
6	24	3.7	74	2	PCT-US95-1 Sequence 94, Applicati	6.68e-02
7	24	3.7	74	2	PCT-US95-1 Sequence 100, Applicat	6.68e-02
8	24	3.7	75	2	PCT-US95-1 Sequence 99, Applicati	6.68e-02
9	24	3.7	81	2	PCT-US95-1 Sequence 92, Applicati	6.68e-02
10	24	3.7	81	2	PCT-US95-1 Sequence 98, Applicati	6.68e-02
11	24	3.7	82	2	PCT-US95-1 Sequence 97, Applicati	6.68e-02
12	23	3.5	66	1	US-08-471- Sequence 144, Applicat	2.15e-01
13	23	3.5	68	1	US-08-471- Sequence 143, Applicat	2.15e-01
14	23	3.5	69	1	US-08-471- Sequence 142, Applicat	2.15e-01
15	23	3.5	74	2	PCT-US95-1 Sequence 100, Applicat	2.15e-01
16	23	3.5	75	2	PCT-US95-1 Sequence 99, Applicati	2.15e-01
17	23	3.5	81	2	PCT-US95-1 Sequence 98, Applicati	2.15e-01
18	23	3.5	82	2	PCT-US95-1 Sequence 97, Applicati	2.15e-01
19	22	3.4	65	1	US-08-471- Sequence 145, Applicat	6.76e-01
20	22	3.4	65	1	US-08-471- Sequence 145, Applicat	6.76e-01

21	22	3.4	66	1	US-08-471- Sequence 144, Applicat	6.76e-01
22	22	3.4	68	1	US-08-471- Sequence 143, Applicat	6.76e-01
23	22	3.4	69	1	US-08-471- Sequence 142, Applicat	6.76e-01
24	22	3.4	242	1	US-08-273- Sequence 1, Applicatio	6.76e-01
25	22	3.4	1995	1	US-08-425- Sequence 3, Applicatio	6.76e-01
26	22	3.4	2256	1	US-08-001- Sequence 1, Applicatio	6.76e-01
27	22	3.4	2256	1	US-07-794- Sequence 1, Applicatio	6.76e-01
28	21	3.2	59	2	PCT-US95-1 Sequence 95, Applicati	2.08e+00
29	21	3.2	59	2	PCT-US95-1 Sequence 95, Applicati	2.08e+00
30	21	3.2	66	2	PCT-US95-1 Sequence 93, Applicati	2.08e+00
31	21	3.2	66	2	PCT-US95-1 Sequence 93, Applicati	2.08e+00
32	21	3.2	108	1	US-08-357- Sequence 2, Applicatio	2.08e+00
33	21	3.2	242	1	US-08-273- Sequence 1, Applicatio	2.08e+00
34	21	3.2	336	1	US-08-248- Sequence 105, Applicat	2.08e+00
35	21	3.2	1441	1	US-08-136- Sequence 18, Applicati	2.08e+00
36	21	3.2	1838	2	PCT-US93-0 Sequence 85, Applicatio	2.08e+00
37	21	3.2	2750	1	US-08-136- Sequence 1, Applicatio	2.08e+00
38	21	3.2	4175	1	US-08-306- Sequence 49, Applicati	2.08e+00
39	21	3.2	4175	2	PCT-US93-0 Sequence 84, Applicati	2.08e+00
40	21	3.2	6063	1	US-08-195- Sequence 4, Applicatio	2.08e+00
41	21	3.2	11558	2	PCT-US93-0 Sequence 23, Applicati	2.08e+00
42	20	3.1	84	1	US-08-133- Sequence 120, Applicat	6.21e+00
43	20	3.1	84	1	US-08-300- Sequence 25, Applicati	6.21e+00
44	20	3.1	84	2	PCT-US95-1 Sequence 25, Applicati	6.21e+00
45	20	3.1	5399	1	US-08-368- Sequence 9, Applicatio	6.21e+00

ALIGNMENTS

RESULT 1  
ID US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.  
AC xxxxxx  
DT  
DE Sequence 14, Application US/08232463  
CC Sequence 14, Application US/08232463  
CC Patent No. 5670367  
CC GENERAL INFORMATION:  
CC APPLICANT: DORNER, F.  
CC APPLICANT: SCHEIFLINGER, F.  
CC APPLICANT: FALKNER, F. G.  
CC TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
CC NUMBER OF SEQUENCES: 52  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Foley & Lardner  
CC STREET: 1800 Diagonal Road, Suite 500  
CC CITY: Alexandria  
CC STATE: VA  
CC COUNTRY: USA  
CC ZIP: 22313-0299  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/232,463  
CC FILING DATE:  
CC CLASSIFICATION: 435  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US/07/935,313  
CC FILING DATE:  
CC APPLICATION NUMBER: EP 91 114 300.6  
CC FILING DATE: 26-AUG-1991  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: BENT, Stephen A.  
CC REGISTRATION NUMBER: 29,768  
CC REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (703)836-9300  
CC TELEFAX: (703)683-4109  
CC TELEX: 899149  
CC INFORMATION FOR SEQ ID NO: 14:  
CC SEQUENCE CHARACTERISTICS:

```
CC LENGTH: 7218 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC CLONE: ptzgpt-F1s
SQ SEQUENCE 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 OTHER.

Query Match      6.0%; Score 39; DB 1; Length 7218;
Best Local Similarity 0.3%; Pred. No. 2.64e-10;
Matches 1; Conservative 203; Mismatches 165; Indels 0; Gaps 0;

Db 1067 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1126
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 13 TCTCGCTGGTACGCGCGCTGCAGAGCCCTCGTGGCACCCTGCTGCTGGTGGG 72

Db 1127 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1186
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 73 CCCTGTCTGGGCACGCTCTACTCAGCCATGGGTACGGTACGGCTCTACCTGGTCTGGAAGA 132

Db 1187 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1246
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 133 GCTGGGAGGCTTCACAGAGAGGCTGTGGTTCCTCCCTGGGCCTCTACACTGGGAGCTGGC 192

Db 1247 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1306
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 193 CCTGAAGTGGGCATGGCCCCCATCTTCTTTGGTGGCCGACAAATGGGCTGGGCTTGGT 252

Db 1307 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1366
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 253 GGATCTCCTGTGTCAGTGGGGCGGCGGAGCCACTACCGTGGCTGTACAGGTGAG 312

Db 1367 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1426
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 313 CCGCTGGCGCGCGCTGCTACTACCCCTACCTGGCCTGGCTGGCTTTCACGACCACT 372

Db 1427 YYYYYYYYG 1435
      : : : : |
QY 373 CAACTACTG 381

RESULT 2
ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
AC xxxxxx
DT
DE Sequence 5, Application US/08238163
CC Sequence 5, Application US/08238163
CC Patent No. 5569830
CC GENERAL INFORMATION:
CC APPLICANT: BENNETT, Alan
CC APPLICANT: LABAVITCH, John M.
CC APPLICANT: POWELL, Ann
CC APPLICANT: STOTZ, Henrik
CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
CC TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
CC NUMBER OF SEQUENCES: 24
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend Kourie and Crew
CC STREET: Steuart Street Tower, One Market Plaza
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: US
CC ZIP: 94105-1493
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/238,163
CC FILING DATE: 03-MAY-1994
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
```

```
CC NAME: Bastian, Kevin L.
CC REGISTRATION NUMBER: 34,774
CC REFERENCE/DOCKET NUMBER: 2307E-540
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 543-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 215 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: misc_feature
CC LOCATION: 1..215
CC OTHER INFORMATION: /standard_name= "Deduced amino acid
CC OTHER INFORMATION: sequence of PGIP from bean."
SQ SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match      5.5%; Score 36; DB 1; Length 215;
Best Local Similarity 13.4%; Pred. No. 1.55e-08;
Matches 28; Conservative 82; Mismatches 98; Indels 1; Gaps 1;

Db 7 SSSVSVRTASCNDKAKKDGNTTSSWTTDCNRTWGVCDTDTTYRVNDSGHNKYSSANY 66
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 504 CGTGACGGGCCACCATCACAAAGCGTGATGGCACCTGCTGGTGCAGCTGCAGTCCCTGGT 445

Db 67 NYGNNVGAATHYTHTNVSGADSKVTDSYNASGTSSSNGGTDGNRSGADSY-GSSKT 125
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 444 GGGCCGGGCACCTACTCTGGCAGCCGCGCTCCCCACGCCAGCCATGGTTGTCCCGCAT 385

Db 126 AMTSRNRGTANNNAVDNRNMGDASVGSDDKNTKKHAKNSADGKVGSKNNGDRNNRYGTGT 185
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 384 ACGCAGTAGTGTAGTGTGGTCGTGAAGCGCAGCCAGGCCAGGTAGGSGTAGAGCAGCGCG 325

Db 186 KSNVSNCGGKNKRDVSSYANNKCCGSSC 214
      : : : : | : : : : | : : :
Cp 324 GCGGCCACCGGCGCTCACCTGGTACCAGGC 296

RESULT 3
ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
AC xxxxxx
DT
DE Sequence 5, Application US/08238163
CC Sequence 5, Application US/08238163
CC Patent No. 5569830
CC GENERAL INFORMATION:
CC APPLICANT: BENNETT, Alan
CC APPLICANT: LABAVITCH, John M.
CC APPLICANT: POWELL, Ann
CC APPLICANT: STOTZ, Henrik
CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
CC TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
CC NUMBER OF SEQUENCES: 24
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend Kourie and Crew
CC STREET: Steuart Street Tower, One Market Plaza
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: US
CC ZIP: 94105-1493
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/238,163
CC FILING DATE: 03-MAY-1994
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
```



```
CC NAME: Bastian, Kevin L.
CC REGISTRATION NUMBER: 34,774
CC REFERENCE/DOCKET NUMBER: 2307E-540
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 543-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 215 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: misc_feature
CC LOCATION: 1..215
CC OTHER INFORMATION: /standard_name= "Deduced amino acid
CC OTHER INFORMATION: sequence of PGIP from bean."
CC SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.
SQ

Query Match 5.4%; Score 35; DB 1; Length 215;
Best Local Similarity 12.4%; Pred. No. 5.93e-08;
Matches 25; Conservative 81; Mismatches 95; Indels 1; Gaps 1;

Db 9 SSVVSRATSCNDKAKKDGNNTSSWTTDCCNRTWGVCDTDTTYRVNNDSGHNKYSSANYNY 68
QY 304 CCAGGTGAGCCCGCTGGCGCCGCCCTGCTCTACCCCTACCTGGCTGGCTGGCTGCAC 363
Db 69 GGNVGAAKTHYTHNVSG-ADSKTVTDSYNASCTSSNGGTDGNSRGADSVSSSKTAM 127
QY 364 GACCACACTCAACTACTGCGTATGGCGGACACCACTGCTGGCTGGCTGGCGGCGGCT 423
Db 128 TSRNRTGKTANNAVDSRNMGDASVSGDKNTKKHAKNSADGKVGSKNNGDRNNRYGTGTS 187
QY 424 GCCAGAGTGAGTCCCGGCCACCAGGACTGCAGTGCACCAAGCAGGTGCCATCAGCT 483
Db 188 NVSNCGGKRDVSSYANNKC 209
QY 484 TGTGATGTGGTGGCGGCTCAGC 505

RESULT 4
ID PCT-US95-11934-94 STANDARD; DNA; UNC; 74 BP.
AC xxxxxx
DT
DE Sequence 94, Application PC/TUS9511934
CC Sequence 94, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
```

```
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 94:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 74 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 74 BP; 3 A; 4 C; 3 G; 1 T; 63 OTHER.
SQ

Query Match 3.8%; Score 25; DB 2; Length 74;
Best Local Similarity 9.9%; Pred. No. 2.03e-02;
Matches 7; Conservative 20; Mismatches 44; Indels 0; Gaps 0;

Db 3 GAGNNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBN 62
QY 432 GAGTGGCCGCCACCAGGACTGCAGCTGCACCAAGCAGGTGCCATCAGCTTGTGATGT 491
Db 63 BNNBNACGCCA 73
QY 492 GGTGGCCGTCA 502

RESULT 5
ID PCT-US95-11934-92 STANDARD; DNA; UNC; 81 BP.
AC xxxxxx
DT
DE Sequence 92, Application PC/TUS9511934
CC Sequence 92, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 92:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 81 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 81 BP; 3 A; 5 C; 6 G; 4 T; 63 OTHER.
SQ

Query Match 3.8%; Score 25; DB 2; Length 81;
Best Local Similarity 9.9%; Pred. No. 2.03e-02;
Matches 7; Conservative 20; Mismatches 44; Indels 0; Gaps 0;
```









```
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-179
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212 790-9090
CC TELEFAX: 212 869-8864/9741
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 143:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 68 bases
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: DNA
SQ SEQUENCE 68 BP; 3 A; 3 C; 5 G; 3 T; 54 OTHER.

Query Match 3.5%; Score 23; DB 1; Length 68;
Best Local Similarity 13.4%; Pred. No. 2.15e-01;
Matches 9; Conservative 18; Mismatches 40; Indels 0; Gaps 0;

Db 1 GGCTAGAVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNV 60
QY 11 GGCTCCGGCTGGTACCGCGGCTGCAGAGCCCTCGTGGCACCGCCCGCCACTGGGTGCTG 70
Db 61 NNACCTG 67
QY 71 GGCCCTG 77

RESULT 14
ID US-08-471-052A-142 STANDARD; DNA; UNC; 69 BP.
AC xxxxxx
DT
DE Sequence 142, Application US/08471052A
CC Sequence 142, Application US/08471052A
CC Patent No. 5625033
CC GENERAL INFORMATION:
CC APPLICANT: Kay, B. K.
CC APPLICANT: Fowkes, D. M.
CC TITLE OF INVENTION: Totally Synthetic Affinity Reagents
CC NUMBER OF SEQUENCES: 166
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036-2711
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/471,052A
CC FILING DATE: 06-JUNE-1995
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-179
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212 790-9090
CC TELEFAX: 212 869-8864/9741
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 142:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 69 bases
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: DNA
SQ SEQUENCE 69 BP; 2 A; 4 C; 6 G; 2 T; 55 OTHER.
```

```
Query Match 3.5%; Score 23; DB 1; Length 69;
Best Local Similarity 13.6%; Pred. No. 2.15e-01;
Matches 9; Conservative 17; Mismatches 40; Indels 0; Gaps 0;

Db 2 GCTCGAGNNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBNBN 61
CP 468 GCTGGTGAGCTGCAGTCCCTCGTGGCGGGGCACTCACTCTGGCAGCCCGCTCCCCCA 409
Db 62 NBCCAG 67
CP 408 CGCCAG 403

RESULT 15
ID PCT-US95-11934-100 STANDARD; DNA; UNC; 74 BP.
AC xxxxxx
DT
DE Sequence 100, Application PC/TUS9511934
CC Sequence 100, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Misrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 100:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 74 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ SEQUENCE 74 BP; 6 A; 6 C; 1 G; 1 T; 60 OTHER.

Query Match 3.5%; Score 23; DB 2; Length 74;
Best Local Similarity 9.0%; Pred. No. 2.15e-01;
Matches 6; Conservative 19; Mismatches 42; Indels 0; Gaps 0;

Db 4 GAVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNV 63
QY 414 GACGGCGGCTGCCAGAGTGAGTGCCCGGCCACCGAGGACTGCAGCTGCACGAGGTG 473
Db 64 NNACCAC 70
QY 474 CCATCAC 480

Search completed: Tue Jul 21 16:08:07 1998
Job time : 41 secs.
```

---

4

\*\*\*\*\*  
M A I N S E Q U E N C E (TM)  
\*\*\*\*\*  
Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm  
Run on: Tue Jul 21 15:51:59 1998; MasPar time 760.56 Seconds  
Tabular output not generated. 1145.078 Million cell updates/sec  
Title: >US-09-047-652A-2  
Description: (1-652) from US09047652A.seq  
Perfect Score: 652  
N.A. Sequence: 1 CCACGGCGAGGGTCTCCGCT.....GTTCTTGGACATGGAATTT 652  
Comp: GGTGCGCTCCAGAGCGGA.....CAAGAACCTGTACCTTAA  
Scoring table: TABLE default  
Gap 6  
Nmatch STD : Dbase 0; Query 0

Searched: 1759237 seqs, 667866413 bases x 2  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: embl-est54  
1:em\_est1 2:em\_est3  
Database: genbank-est106  
3:gb\_est1 4:gb\_est10 5:gb\_est11 6:gb\_est12 7:gb\_est13  
8:gb\_est14 9:gb\_est15 10:gb\_est16 11:gb\_est17  
12:gb\_est18 13:gb\_est19 14:gb\_est2 15:gb\_est20  
16:gb\_est21 17:gb\_est22 18:gb\_est3 19:gb\_est4 20:gb\_est5  
21:gb\_est6 22:gb\_est7 23:gb\_est8 24:gb\_est9 25:gb\_gss  
26:gb\_sts  
Statistics: Mean 10.451; Variance 2.423; scale 4.313  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES									
Result No.	Score	Query Match %	Length	DB	ID	Description	Pred. No.		
C 1	498	76.4	546	15	AA724486	ah99hl1.s1 Soares NFL	0.00e+00		
C 2	493	75.6	572	9	AA587126	nn70bl2.s1 NCI_CGAP_La	0.00e+00		
C 3	463	71.0	567	13	AA775735	zf31f04.s1 Soares feta	0.00e+00		
C 4	459	70.4	563	23	AA161033	zo58e04.s1 Stratagene	0.00e+00		
C 5	458	70.2	610	12	AA069479	zm14hl1.s1 Stratagene	0.00e+00		
C 6	452	69.3	568	10	AA595715	n146e08.s1 NCI_CGAP_Lu	0.00e+00		
C 7	446	68.4	591	11	AA643068	nr95f04.s1 NCI_CGAP_Pr	0.00e+00		
C 8	442	67.8	556	8	AA210894	zr90g06.s1 NCI_CGAP_GC	0.00e+00		
C 9	437	67.0	541	11	AA479680	zu43a07.s1 Soares ovar	0.00e+00		
C 10	425	65.2	541	15	AA748434	ny01bl0.s1 NCI_CGAP_GC	0.00e+00		
C 11	424	65.0	506	7	AA456315	aal3gl0.s1 Soares NhHM	0.00e+00		
C 12	414	63.5	551	5	AA036726	zk30b07.s1 Soares preg	0.00e+00		
C 13	405	62.1	505	16	AA838533	oe39a11.s1 NCI_CGAP_Pr	0.00e+00		

C 14	399	61.2	543	15	AA805072	ob86c01.s1 NCI_CGAP_GC	0.00e+00
C 15	394	60.4	507	13	AA768741	ob22g01.s1 NCI_CGAP_K1	0.00e+00
C 16	392	60.1	558	11	AA479803	zu43a07.r1 Soares ovar	0.00e+00
C 17	390	59.8	457	7	AA455945	aal6e07.s1 Soares NHM	0.00e+00
C 18	386	59.2	505	16	AA768733	oc86h05.s1 NCI_CGAP_GC	0.00e+00
C 19	386	59.2	516	9	AA595368	nc36d03.s1 NCI_CGAP_Pr	0.00e+00
C 20	376	57.7	540	12	AA654418	nt03h01.s1 NCI_CGAP_Ly	0.00e+00
C 21	375	57.5	465	11	AA653746	ns63d09.s1 NCI_CGAP_Pr	0.00e+00
C 22	373	57.2	521	9	AA559042	nl11e02.s1 NCI_CGAP_Pr	0.00e+00
C 23	364	55.8	504	15	AA805330	oc15c05.s1 NCI_CGAP_GC	0.00e+00
C 24	360	55.2	490	15	AA857072	oe33d06.s1 NCI_CGAP_Pr	0.00e+00
C 25	360	55.2	498	16	AA808022	oc38e02.s1 NCI_CGAP_GC	0.00e+00
C 26	356	54.6	481	15	AA834034	of25e01.s1 NCI_CGAP_K1	0.00e+00
C 27	346	53.1	541	9	AA535901	nf94b05.s1 NCI_CGAP_Co	0.00e+00
C 28	345	52.9	444	22	AA100452	zl81c10.s1 Stratagene	0.00e+00
C 29	342	52.5	401	17	AA904236	oe73g05.s1 NCI_CGAP_Lu	0.00e+00
C 30	341	52.3	419	22	AA099285	zl77h01.s1 Stratagene	0.00e+00
C 31	337	51.7	424	23	AA158134	zo43hl1.s1 Stratagene	0.00e+00
C 32	322	49.4	495	23	AA158833	zo58e04.r1 Stratagene	0.00e+00
C 33	315	48.3	402	19	N23981	yx71g01.s1 Homo sapien	0.00e+00
C 34	314	48.2	323	4	AA313263	EST185182 Colon carcin	0.00e+00
C 35	314	48.2	416	22	W44958	zc20g05.s1 Soares sene	0.00e+00
C 36	311	47.7	446	16	AA837816	oe39h02.s1 NCI_CGAP_Pr	0.00e+00
C 37	310	47.5	541	22	W25523	zb90b12.r1 Soares sene	0.00e+00
C 38	308	47.2	471	22	AA056606	zl66c07.s1 Stratagene	0.00e+00
C 39	308	47.2	578	8	AA210893	zr90g06.r1 NCI_CGAP_GC	0.00e+00
C 40	307	47.1	462	20	N63826	za26e09.s1 Homo sapien	0.00e+00
C 41	307	47.1	468	22	AA037186	zc51c06.r1 Soares sene	0.00e+00
C 42	306	46.9	374	21	W80770	zd90c02.s1 Soares feta	0.00e+00
C 43	297	45.6	423	22	W70148	zd52g04.s1 Soares feta	0.00e+00
C 44	293	44.9	460	20	N94184	za26e09.r1 Homo sapien	0.00e+00
C 45	286	43.9	430	19	N31882	yx71g01.r1 Homo sapien	0.00e+00

ALIGNMENTS

RESULT 1  
LOCUS ah99hl1.s1 Soares NFL T GBC S1 Homo sapiens cDNA clone 1327269 3'  
DEFINITION similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR (HUMAN);, mRNA sequence.  
ACCESSION AA724486 546 bp mRNA EST 11-FEB-1998  
NID ah99hl1.s1  
KEYWORDS g2742193  
SOURCE EST.  
ORGANISM human.  
REFERENCE Homo sapiens  
AUTHORS Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
TITLE Primates; Catarrhini; Homnidae; Homo.  
JOURNAL 1 (bases 1 to 546)  
COMMENT NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)

Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 472.  
Location/Qualifiers  
1. 546  
/organism="Homo sapiens"  
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with  
a modified polylinker; Site\_1: Not 1; Site\_2: Eco RI;  
Equal amounts of plasmid DNA from three normalized  
libraries (fetal lung NBHL19W, testis NHT, and B-cell  
NCI\_CGAP\_GCB1) were mixed, and ss circles were made in  
vitro. Following HAP purification, this DNA was used as  
tracer in a subtractive hybridization reaction. The driver  
was PCR-amplified cDNAs from pools of 5,000 clones made  
from the same 3 libraries. The pools consisted of

I.M.A.G.E. clones 297480-302087, 682632-687239,  
726408-728711, and 729096-731399. Subtraction by Bento  
Soares and M. Fatima Bonaudo."  
/db\_xref="taxon:9606"  
/clone="1327269"  
/clone\_lib="Soares NFL T GBC S1"  
/lab\_host="DH10B"  
BASE COUNT 120 a 166 c 166 g 91 t 3 others  
ORIGIN

Query Match 76.4%; Score 498; DB 15; Length 546;  
Best Local Similarity 98.4%; Pred. No. 0.00e+00;  
Matches 503; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Db 36 AAATTCCATGTTCCAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 95  
|||||  
Cp 652 AAATTCCATGTTCCAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 593  
Db 96 GTGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGTGACCCCTGGGCCAAGSCCTGAC 155  
|||||  
Cp 592 GTGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGTGACCCCTGGGCCAAGSCCTGAC 533  
Db 156 AGACTAGCAGGCCCTCAGTGGTGCATGAAGCGTGCAGCGCCACCATCACAAGCGTGATGGC 215  
|||||  
Cp 532 AGACTAGCAGGCCCTCAGTGGTGCATGAAGCGTGCAGCGCCACCATCACAAGCGTGATGGC 473  
Db 216 ACCTGCTGGTCANGCTGCAGTCCCTGGTGGCGGGGCGGACTCACTCTGGCAGCGCGCTCC 275  
|||||  
Cp 472 ACCTGCTGGTCANGCTGCAGTCCCTGGTGGCGGGGCGGACTCACTCTGGCAGCGCGCTCC 413  
Db 276 CCACGCCAGCCATGGTTGCCGCCATACGCAGTAGTGTGAGTGTGGTCTGTAAGGCCAG 335  
|||||  
Cp 412 CCACGCCAGCCATGGTTGCCGCCATACGCAGTAGTGTGAGTGTGGTCTGTAAGGCCAG 353  
Db 336 CCAGGCCAGGTAGGGGTAGAGCAGGGCGGGCGGCCAGCGGGCTCACCTGGTACCAGGCCAC 395  
|||||  
Cp 352 CCAGGCCAGGTAGGGGTAGAGCAGGGCGGGCGGCCAGCGGGCTCACCTGGTACCAGGCCAC 293  
Db 396 GGTAAGTGTGNCNCGCCGCTGACAGCAGGAGATCCACCAAGGCCAGGCCATTG 455  
|||||  
Cp 292 GGTAAGTGTGNCNCGCCGCTGACAGCAGGAGATCCACCAAGGCCAGGCCATTG 233  
Db 456 TCGGGCACCAAGAAGATGGGGGCCATGCCAGTTTTCAGGGCCAGCTGCCAGTGTAGAG 515  
|||||  
Cp 232 TCGGGCACCAAGAAGATGGGGGCCATGCCAGTTTTCAGGGCCAGCTGCCAGTGTAGAG 173  
Db 516 GCCCAGGGGAACACAGCCTTCTCTGTGAAG 546  
|||||  
Cp 172 GCCCAGGGGAACACAGCCTTCTCTGTGAAG 142

RESULT 2  
LOCUS AA587126 572 bp mRNA EST 12-SEP-1997  
DEFINITION nn70b12.s1 NCI\_CGAP\_Lar1 Homo sapiens CDNA clone IMAGE:1089215  
similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR  
(HUMAN);, mRNA sequence.  
ACCESSION AA587126  
NID g2397940  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 572)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert\_Strausberg@nih.gov  
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.

Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: Stratagene, Inc.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 462.  
FEATURES  
Location/Qualifiers  
1..572

/organism="Homo sapiens"  
/note="Organ: larynx; Vector: Bluescript SK-; Site\_1:  
EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer:  
Oligo dT. Larynx. 5' adaptor sequence: 5' GAATTCGGCAGG  
3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'  
Average insert size: 0.9 kb."  
/db\_xref="taxon:9606"  
/clone="IMAGE:1089215"  
/clone\_lib="NCI\_CGAP\_Lar1"  
/tissue\_type="larynx"  
/lab\_host="SOLR (kanamycin resistant)"  
BASE COUNT 132 a 171 c 173 g 96 t  
ORIGIN

Query Match 75.6%; Score 493; DB 9; Length 572;  
Best Local Similarity 96.5%; Pred. No. 0.00e+00;  
Matches 522; Conservative 0; Mismatches 17; Indels 2; Gaps 2;

Db 34 AAATTCCATGTTCCAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 93  
|||||  
Cp 652 AAATTCCATGTTCCAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 593  
Db 94 GTGGGGCTCAGGTGGGGCAACCTCTGAAGCTCTGTGACCCCTGGGCCAAGGCCCTGAC 153  
|||||  
Cp 592 GTGGGGCTCAGGTGGGGCAACCTCTGAAGCTCTGTGACCCCTGGGCCAAGGCCCTGAC 533  
Db 154 AGACTAGCAGGCCCTCAGTGGTGCATGAAAGCGTGACGGCCACCATCACAAGCGTGATGGC 213  
|||||  
Cp 532 AGACTAGCAGGCCCTCAGTGGTGCATGAAAGCGTGACGGCCACCATCACAAGCGTGATGGC 473  
Db 214 ACCTGCTGGTGCAGTGCAGTCCCTGGTGGGGCGGGCAGTCACTCTGGCAGCCCGCTCC 273  
|||||  
Cp 472 ACCTGCTGGTGCAGTGCAGTCCCTGGTGGGGCGGGCAGTCACTCTGGCAGCCCGCTCC 413  
Db 274 CCCATGCCAGCCATGGTTGTCCCGCCATACGCAGTAGTTGAGTGTGGTCGCGAAGGCCAG 333  
|||||  
Cp 412 CCCACGCCAGCCATGGTTGTCCCGCCATACGCAGTAGTTGAGTGTGGTCGTAAGGCCAG 353  
Db 334 CCAGGCCAGGTAGGGGTAGAGCAGGGCGGGCCAGCGGGCTCACCTGGTACCAGGCCAC 393  
|||||  
Cp 352 CCAGGCCAGGTAGGGGTAGAGCAGGGCGGGCCAGCGGGCTCACCTGGTACCAGGCCAC 293  
Db 394 GGTAAGTGTGGACCGCCCGCCCTGACCTGACCAAGCAGGAGATCCACCAAGSCCAGCCATTG 453  
|||||  
Cp 292 GGTAAGTGTGGACCGCCCGCCCTGACCTGACCAAGCAGGAGATCCACCAAGSCCAGCCATTG 233  
Db 454 TCGGGCACCAAGAAGATGGGGGAGCCCATGCCAGTTTCAGGGCCAGCTGCACAGTGTAGAG 513  
|||||  
Cp 232 TCGGGCACCAAGAAGATGGGGGAGCCATGCCAGTTTCAGGGCCAGCTGCACAGTGTAGAG 173  
Db 514 GCCCAGGGGAAC-ACAGCCTTCTCTGTGAAGACTCCCAGCTCTTTC-AGACCAGGTATGA 571  
|||||  
Cp 172 GCCCAGGGGAACACAGCCTTCTCTGTGAAGCCTCCCAGCTCTTTCAGACCAGGTATGA 113  
Db 572 G 572  
Cp 112 G 112  
RESULT 3



LOCUS AA775735 567 bp mRNA EST 05-FEB-1998  
DEFINITION zf31f04.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone  
378559 3' similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE  
RECEPTOR (HUMAN);, mRNA sequence.  
ACCESSION AA775735  
NID 92835069  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 567)  
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,  
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,  
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,  
Theising,B., White,Y., Wyllie,T., Waterston,R. and Wilson,R.  
TITLE WashU-NCI human EST Project  
JOURNAL Unpublished (1997)  
COMMENT  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 487.  
FEATURES  
source  
1. .567  
/organism="Homo sapiens"  
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a  
modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer [5'  
TGTTACCAATCTGAAGTGGGAGCGCGGCATCTTTTCTTTTCTTTT 3'],  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified pT7T3 vector  
(Pharmacia). Library went through one round of  
normalization to a Cot = 5. Library constructed by  
M.Fatima Bonaldo. This library was constructed from the  
same fetus as the fetal lung library, Soares fetal lung  
NbHL19W."  
/db\_xref="GDB:1286815"  
/db\_xref="taxon:9606"  
/clone="378559"  
/clone\_lib="Soares fetal heart NbHH19W"  
/sex="unknown"  
/dev\_stage="19 weeks"  
/lab\_host="DH10B (ampicillin resistant)"  
BASE COUNT 126 a 167 c 175 g 98 t 1 others  
ORIGIN  
Query Match 71.0%; Score 463; DB 13; Length 567;  
Best Local Similarity 96.3%; Pred. NO. 0.00e+00;  
Matches 522; Conservative 0; Mismatches 12; Indels 8; Gaps 8;  
Db 30 AAATCCATGTTCCAGAACATGCTCTAAGCATGCGAAGACACAGGACACTGCTCCCGG 89  
|||||  
Cp 652 AAATCCATGTTCCAGAACATGCTCTAAGCATGCGAAGACACAGGACACTGCTCCCGG 593  
Db 90 GTGGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGTGACCCCTGGGCCAAGCGCTGTAC 149  
|||||  
Cp 592 GTGGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGTGACCCCTGGGCCAAGCGCTGTAC 533  
Db 150 AGACTAGCAGGCCCGGCTCATGAAGCGGTGACGGCCACCACATCACAAAGCGTGATGGC 209  
|||||  
Cp 532 AGACTAGCAGGCCCGGCTCATGAAGCGGTGACGGCCACCACATCACAAAGCGTGATGGC 473  
Db 210 ACCTGCTGGTG-AGCTGCAGTCCCTGGTGGGGCGGGGCACTCACTCTGGCAGCGCGCTCC 268  
|||||  
Cp 472 ACCTGCTGGTGACAGCTGCAGTCCCTGGTGGGGCGGGCACTCACTCTGGCAGCGCGCTCC 413

Db 269 CCCACGCCAGCATGGTTGTCCCGCCATACGCAGTAGTTAGTGTGCTGTAAGGCCAG 328  
|||||  
Cp 412 CCCACGCCAGCATGGTTGTCCCGCCATACGCAGTAGTTAGTGTGCTGTAAGGCCAG 353  
Db 329 CCAGGCTAGGTAGGGGTAGAGCAGGCGGGGCCAGCGGGCTCACCTGGTACCAGGCCAC 388  
|||||  
Cp 352 CCAGGCCAGGTAGGGGTAGAGCAGGCGGGGCCAGCGGGCTCACCTGGTACCAGGCCAC 293  
Db 389 GGTAGTGTGACTGCCCGCCCCACTGACCAGCAGGAGATCCACCAAGGCCAGCCCCATTG 448  
|||||  
Cp 292 GGTAGTGGCTGCCCGCCCCACTGACCAGCAGGAGATCCACCAAGGCCAGCCCCATTG 233  
Db 449 TCGGGCAC-ANAGAAGATGGGGGCCATGCC-AGTTCAGGGCCAGCTGCCTAGTGTAGAG 506  
|||||  
Cp 232 TCGGGCACCAAGAAGATGGGGGCCATGCCAGTTCAGGGCCAGCTGCCAGTGTAGAG 173  
Db 507 GCC-AGAGGAAC-ACAGCCTTCTCTGTGAAGCCTCC-AGCTCTTTC-AGAC-AGGTAGGA 561  
|||||  
Cp 172 GCCCAGGGGAACACAGCCTTCTCTGTGAAGCCTCCAGCTCTTCCACAGCAGGTAGGA 113  
Db 562 GC 563  
||  
Cp 112 GC 111  
RESULT 4  
LOCUS AA161033 563 bp mRNA EST 16-DEC-1996  
DEFINITION z058e04.s1 Stratagene pancreas (#937208) Homo sapiens cDNA clone  
591102 3' similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE-  
RECEPTOR (HUMAN);, mRNA sequence.  
ACCESSION AA161033  
NID g1735356  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 563)  
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,  
Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,  
Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.  
TITLE WashU-Merck EST Project  
JOURNAL Unpublished (1995)  
COMMENT  
Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40M13 fwd. from Amersham  
High quality sequence stop: 425.  
FEATURES  
Location/Qualifiers  
source  
1. .563  
/organism="Homo sapiens"  
/note="Organ: pancreas; Vector: pBluescript SK-; Site\_1:  
EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer:  
Oligo dT. Pancreatic adenocarcinoma cell line. Average  
insert size: 1.0 kb; Uni-ZAP XR vector; -5' adaptor  
sequence: 5' GAATTCGCGCAGGAG 3' -3' adaptor sequence: 5'  
CTCGAGTCTTTTCTTTTCTTTT 3,"  
/db\_xref="taxon:9606"  
/clone="591102"  
/clone\_lib="Stratagene pancreas (#937208)"  
/lab\_host="SOLR cells (kanamycin resistant)"  
complement(<1..>563)  
BASE COUNT 127 a 168 c 169 g 98 t 1 others  
ORIGIN

Query Match 70.4%; Score 459; DB 23; Length 563;  
Best Local Similarity 96.2%; Pred. No. 0.00e+00;  
Matches 513; Conservative 0; Mismatches 13; Indels 7; Gaps 7;

Db 35 AAATTCCATGTTCCAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 94  
|||||  
Cp 652 AAATTCCATGTTCCAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 593

Db 95 GTGGGGGCTCAGGTGGGCAACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGCCCATGAC 154  
|||||  
Cp 592 GTGGGGGCTCAGGTGGGCGCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGCCCTGAC 533

Db 155 AGACTAGCAGGCCCGCAGTGGTCATGAAAGCGTGACGGCCACCATCACAAAGCGTGATGGC 214  
|||||  
Cp 532 AGACTAGCAGGCCCGCAGTGGTCATGAAAGCGTGACGGCCACCATCACAAAGCGTGATGGC 473

Db 215 ACCTGCTGGTN-AGCTGCAGTCCCTGCTGGTGGCCGCGGCACACTCACTCTGGCAGCGCCGTCC 273  
|||||  
Cp 472 ACCTGCTGGTGCACTGAGTCCCTGCTGGTGGCCGCGGCACACTCACTCTGGCAGCGCCGTCC 413

Db 274 CCCATGCCAGCCATGTTGTCGCCCATACGCAGTAGTTGAGTGTGGTCCGCGAAGGCCAG 333  
|||||  
Cp 412 CCCAGCCAGCCATGTTGTCGCCCATACGCAGTAGTTGAGTGTGGTCTGCGAAGGCCAG 353

Db 334 CCAGGCCAGGTAAGGGGTAGAGCAGGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 393  
|||||  
Cp 352 CCAGGCCAGGTA-GGGGTAGAGCAGGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 296

Db 394 CACGGTAGTTGGC-CCGCGCGCCCACTGACCAAGCAGGAGATCCACCAAGGCCAGCCCAT 452  
|||||  
Cp 295 CACGGTAGTTGGCTGCCGCGCCCACTGACCAAGCAGGAGATCCACCAAGGCCAGCCCAT 236

Db 453 TTGTCGGGCACCAAGAGATGGGGGCCCATGCAAGTTTCAGGGCCAGTGCC-AGTGTA 511  
|||||  
Cp 235 TTGTCGGGCACCAAGAGATGGGGGCCCATGCCCAAGTTTCAGGGCCAGTGCCAGTGTA 176

Db 512 GAGGCCAGGGGAACCAACAGCCTTCTCTGTGAAGCTTCC-AGCTCTTTCCAGA 563  
|||||  
Cp 175 GAGGCCAGGGGAACCAACAGCCTTCTCTGTGAAGCCTCCAGCTCTTTCCAGA 123

RESULT 5  
LOCUS AA069479 610 bp mRNA EST 23-DEC-1997  
DEFINITION zml4h11.s1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 525669 3' similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR (HUMAN);, mRNA sequence.

ACCESSION AA069479  
NID g1577104  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 610)  
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.  
WashU-Merck EST Project  
Unpublished (1995)

Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
WARNING: There is evidence that suggests that the 384-well parent plate of this clone contains both human and mouse derived clones. Thus, the origin of this clone is uncertain. This caution should be kept in mind should you use this clone.

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert length: 951 Std Error: 0.00  
Seq primer: -40M13 fwd. from Amersham  
High quality sequence stop: 118.

FEATURES  
source  
location/Qualifiers  
1. .610  
/organism="Homo sapiens"  
/note="Organ: pancreas; Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Pancreatic adenocarcinoma cell line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCACAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3,"  
/db\_xref="GDB:3917098"  
/db\_xref="taxon:9606"  
/clone="525669"  
/clone\_lib="Stratagene pancreas (#937208)"  
/lab\_host="SOLR cells (kanamycin resistant)"

BASE COUNT 133 a 177 c 175 g 103 t 22 others  
ORIGIN

Query Match 70.2%; Score 458; DB 12; Length 610;  
Best Local Similarity 92.3%; Pred. No. 0.00e+00;  
Matches 527; Conservative 0; Mismatches 35; Indels 9; Gaps 8;

Db 33 AAATTCCATGTTCCAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 92  
|||||  
Cp 652 AAATTCCATGTTCCAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 593

Db 93 GTGGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGSCCAAGGCC-TGAC 151  
|||||  
Cp 592 GTGGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGSCCAAGGCCCTGAC 533

Db 152 AGACTANAGGCCCGCAGTGGTCATGAAGCGTGACGGCCACCATCACATCACAAAGCGTGATGGC 211  
|||||  
Cp 532 AGACTAGAGGCCCGCAGTGGTCATGAAGCGTGACGGCCACCATCACATCACAAAGCGTGATGGC 473

Db 212 ACCTGCTGGT-CAGCTGCAGTCCCTGCTGGGCGCGGCGCAGTCACTCTGGCAGNCGCGTCC 270  
|||||  
Cp 472 ACCTGCTGGTGCACTGCACTCCCTGCTGGGCGCGGCGCAGTCACTCTGGCAGNCGCGTCC 413

Db 271 CCCACGCCAGCCATNGTTGTCCCGCCNFACTCAGTAGTTGAGTNTGTCGCGAAGGCCAG 330  
|||||  
Cp 412 CCCACGCCAGCCATNGTTGTCCCGCCATACGACAGTAGTTGAGTGTGTCGTGAAGGCCAG 353

Db 331 CCAGGCCAGGTATGGGTAGAGCAGGCGGGCGGCCAGCGGGCTCACCTGGTACCAGGCCAC 390  
|||||  
Cp 352 CCAGGCCAGGTATGGGTAGAGCAGGCGGGCGGCCAGCGGGCTCACCTGGTACCAGGCCAC 293

Db 391 GGTATTGTC--CGCGCGCCCGCCACTNACGAGNANNAGGATTTCACCAAGSCCCAGCCCATTT 448  
|||||  
Cp 292 GGTAGTGGCTGCCGCGCCCGCCACTGACCAAGCAGGAG-ATCCACCAAGSCCCAGCCCATTT 234

Db 449 GTCGGGGACCAAGAGATNGGGGGCGCATGCC-AGTTCAGGGC-AGNTGCCNAGTGTAGA 506  
|||||  
Cp 233 GTCGGGCACCAAGAGATNGGGGGCGCATGCCAGTTCAGGGCCAGTGTGCCAGTGTAGA 174

Db 507 GGCCCAAGGTGAACACANGCTTCTCTGTGNAGC-TCCCAG-TNTTTCAGACCAAGTNNNG 564  
|||||  
Cp 173 GGCCCAAGGGGAACCAAGCCTTCTCTGTGAAGCCTCCAGCTCTTTCCAGACCAAGGTAGG 114

Db 565 AGCCGTNNCCATGGCTANGTAGAGCTTCC 595  
|||||  
Cp 113 AGCCGTACCCCATGGCTGAGTAGAGCGTGCC 83

RESULT 6  
LOCUS AA595715 568 bp mRNA EST 25-SEP-1997  
DEFINITION ni46e08.s1 NCI\_CGAP\_Lu1 Homo sapiens cDNA clone IMAGE:979910 similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR (HUMAN);, mRNA sequence.  
ACCESSION AA595715

NID 92411065  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 568)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: Stratagene, Inc., David B. Krizman,  
Ph.D.  
cDNA Library Arraying: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Insert Length: 713 Std Error: 0.00  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 265.  
FEATURES Location/Qualifiers  
source 1..568  
/organism="Homo sapiens"  
/note="Organ: lung; Vector: Bluescript SK-; Site\_1: EcoRI;  
Site\_2: XhoI; Cloned unidirectionally. Primer: Oligo dT.  
Bulk lung tumor. 5' adaptor sequence: 5' GAATTCGGCAGG 3'  
3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'  
Average insert size: 1.1 kb."  
/db\_xref="taxon:9606"  
/clone="IMAGE:979910"  
/clone\_lib="NCI\_CGAP\_Lul"  
/tissue\_type="lung tumor"  
/lab\_host="SOLR (kanamycin resistant)"  
BASE COUNT 124 a 164 c 165 g 113 t 2 others  
ORIGIN  
Query Match 69.3%; Score 452; DB 10; Length 568;  
Best Local Similarity 96.9%; Pred. No. 0.00e+00;  
Matches 496; Conservative 0; Mismatches 10; Indels 6; Gaps 6;  
Db 53 AAATTCATGTTCCAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 112  
|||||  
Cp 652 AAATTCATGTTCCAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 593  
|||||  
Db 113 GTGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCCCTGAC 172  
|||||  
Cp 592 GTGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCCCTGAC 533  
|||||  
Db 173 AGACTAGCAGGCCCGGCTCATGAAAGCGTGACGGCCACCACATCACACGCGTGATGGC 232  
|||||  
Cp 532 AGACTAGCAGGCCCGGCTCATGAAAGCGTGACGGCCACCACATCACACGCGTGATGGC 473  
|||||  
Db 233 ACCTGCTGGTGTAGTGCAGTCCCTGGTGGCGGGGCACTCACTCTGGCAGCGCCGTCC 292  
|||||  
Cp 472 ACCTGCTGGTGCAGTGCAGTCCCTGGTGGCGGGGCACTCACTCTGGCAGCGCCGTCC 413  
|||||  
Db 293 CCCACGCCAGCCATGGTGTCCCGCCATACGCAGTAGTTGAGTGTGGTCGTGAAGGCCAG 352  
|||||  
Cp 412 CCCACGCCAGCCATGGTGTCCCGCCATACGCAGTAGTTGAGTGTGGTCGTGAAGGCCAG 353  
|||||  
Db 353 CCAGGCCAGGTAGGGGTAGAGCAGGCGGGCCCGGCGGCTCACCTGGTACCAGGCCAC 412  
|||||  
Cp 352 CCAGGCCAGGTAGGGGTAGAGCAGGCGGGCCCGGCGGCTCACCTGGTACCAGGCCAC 293  
|||||  
Db 413 GGTACTGGNN-CCGCGCGCCCACTGACCCAGCAGGAGATCCACCAAGGCCCAAGCCCATTTG 471

Cp 292 GGTAGTGGTCCCGCCGCCCCACTGACCAGCAGGAGATCCACCAAGGCCCAAGCCCATTTG 233  
|||||  
Db 472 TCGGGCACCAAGAAGATGGGGGGCCATGCAC-GTTCTAGGGCAGCTGCCAA-TGTAGAG 529  
|||||  
Cp 232 TCGGGCACCAAGAAGATGGGGGGCCATGCCAGTTTCAGGGCCAGCTGCCCAGTGTAGAG 173  
|||||  
Db 530 GCCAAGGG-AAC-ACAGC-TTCTCTGTGAAGC 558  
|||||  
Cp 172 GCCCAGGGGAACCCACAGCCCTTCTCTGTGAAGC 141  
|||||  
RESULT 7  
LOCUS AA643068 591 bp mRNA EST 27-OCT-1997  
DEFINITION nr95f04.s1 NCI\_CGAP\_Pr25 Homo sapiens cDNA clone IMAGE:1175743  
similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR  
(HUMAN);, mRNA sequence.  
ACCESSION AA643068  
NID g2568286  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 591)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: Suzanne L. Topalian, M.D., Robert K. Bright,  
Ph.D.  
cDNA Library Preparation: Stratagene, Inc.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 337.  
FEATURES Location/Qualifiers  
source 1..591  
/organism="Homo sapiens"  
/note="Organ: prostate; Vector: Bluescript SK-; Site\_1:  
EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer:  
Oligo dT. Normal prostate epithelial cell line (HPV  
immortalized). 5' adaptor sequence: 5' GAATTCGGCAGG 3'  
3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'  
Average insert size: 1.1 kb."  
/db\_xref="taxon:9606"  
/clone="IMAGE:1175743"  
/clone\_lib="NCI\_CGAP\_Pr25"  
/tissue\_type="epithelium (cell line)"  
/lab\_host="SOLR (kanamycin resistant)"  
BASE COUNT 131 a 179 c 181 g 97 t 3 others  
ORIGIN  
Query Match 68.4%; Score 446; DB 11; Length 591;  
Best Local Similarity 95.4%; Pred. No. 0.00e+00;  
Matches 544; Conservative 0; Mismatches 11; Indels 15; Gaps 15;  
Db 28 AAATTCATGTTCCAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 87  
|||||  
Cp 652 AAATTCATGTTCCAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 593  
|||||  
Db 88 GTGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCAAGGCCCTGAC 147  
|||||  
Cp 592 GTGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCAAGGCCCTGAC 533  
|||||



Db 148 AGACTAGCAGGCCAGTGGTTCATGAAAGCGTGAACGGACGGCCACCACATCACAAAGCGTGATGGC 207  
|||||  
Cp 532 AGACTAGCAGGCCAGTGGTTCATGAAAGCGTGAACGGACGGCCACCACATCACAAAGCGTGATGGC 473  
  
Db 208 ACCTGCTGGTG-AGCTGCAGTCCCTGCTGGGCGGGCCAGCTCACTCTGGCAGCCGC-GTCC 265  
|||||  
Cp 472 ACCTGCTGGTGAGCTGCAGTCCCTGCTGGGCGGGCCAGCTCACTCTGGCAGCCGCCTCC 413  
  
Db 266 CCCACGCCAGCCATGGTTGTCCCGCCCATACGACAGTAGTTGAGTGTGTCGCGAAGGCCAG 325  
|||||  
Cp 412 CCCACGCCAGCCATGGTTGTCCCGCCCATACGACAGTAGTTGAGTGTGTCGCGAAGGCCAG 353  
  
Db 326 CCAGGCCAGGTAGGGGTAGACAGGGCGGGCCGACGGGCTCACCTGGTACCAGGCCAC 385  
|||||  
Cp 352 CCAGGCCAGGTAGGGGTAGACAGGGCGGGCCGACGGGCTCACCTGGTACCAGGCCAC 293  
  
Db 386 GGTACTGGNCNC-GCCGCCCCACTGACACGAGGAGATCCACCAGGCCAGCCCATTTG 444  
|||||  
Cp 292 GGTAGTGGCTGCCGCGCCCCACTGACACGAGGAGATCCACCAGGCCAGCCCATTTG 233  
  
Db 445 TCGGGCAC-AAAGAGATGGGGGCCCATGCC-AGTTCAGGGCCAGCTGCC-AGTGTAGAN 501  
|||||  
Cp 232 TCGGGCACCAAAGAGATGGGGGCCCATGCCAGTTCAGGGCCAGCTGCCCATTTG 173  
  
Db 502 GCC-AGGGGAAC-ACAGC-TTCTCTGTGAAGCTCC-AGCTCTTTC-AGAC-AGGTAGGA 555  
|||  
Cp 172 GCCAGGGGAACACAGCTTCTCTGTGAAGCTCCAGCTCTTTCCAGACCAGGTAGGA 113  
  
Db 556 GC-GTACCC-ATG-CTGAGTAGACGTTGCC 582  
|||  
Cp 112 GCCGTACCCCATGGCTGAGTAGACGTTGCC 83  
  
RESULT 8  
LOCUS AA210894 556 bp mRNA EST 13-AUG-1997  
DEFINITION zr90q06.s1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:683002 3'  
similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR  
(HUMAN);, mRNA sequence.  
ACCESSION AA210894  
NID g1809540  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;  
Homo.  
REFERENCE 1 (bases 1 to 556)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT

Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert\_Strausberg@nih.gov  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -41ml3 fwd. Et from Amersham  
High quality sequence stop: 428.  
Location/Qualifiers  
1. .556  
/organism="Homo sapiens"  
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was prepared from human tonsillar cells enriched for  
germinal center B cells by flow sorting (CD20+, IgD-),  
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was  
primed with a Not I - oligo(dT) primer  
[5'-TGTTACCAATCTGAAGTGGGAGCGCGCCCTCATTTTTTTTTTTTTTT-  
3']. Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT7T3 vector. Library

went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."  
/db\_xref="taxon:9606"  
/clone\_lib="NCI\_CGAP\_GCB1"  
/tissue\_type="germinal center B cell"  
/lab\_host="DH10B"  
complement(<1..>556)  
/db\_xref="GDB:5586322"  
BASE COUNT 122 a 160 c 164 g 104 t 6 others  
ORIGIN  
  
Query Match 67.8%; Score 442; DB 8; Length 556;  
Best Local Similarity 96.9%; Pred. No. 0.00e+00;  
Matches 473; Conservative 0; Mismatches 11; Indels 4; Gaps 4;  
  
Db 43 AAATTCCATGTTCCAAGAACATGCTCTAAGCATGCAGAAACACACAGGACACTGCTCCCG 102  
|||||  
Cp 652 AAATTCCATGTTCCAAGAACATGCTCTAAGCATGCAGAAACACACAGGACACTGCTCCCG 593  
  
Db 103 GTGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGCCAAAGCCNTGAC 162  
|||||  
Cp 592 GTGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGCCAAAGCCCTGAC 533  
  
Db 163 AGACTAGCAGGCCACAGTGGTCAATGAAGCGTGAACGCGGACCATCATCACAAAGCGTGATGGC 222  
|||||  
Cp 532 AGACTAGCAGGCCACAGTGGTCAATGAAGCGTGAACGCGGACCATCATCACAAAGCGTGATGGC 473  
  
Db 223 ACCTGCTGGTN-AGCTGCAGTCCCTGGTGGCGGCGGCACTCACTCTGGCAGCCCGCTCC 281  
|||||  
Cp 472 ACCTGCTGGTGCAGCTGCAGTCCCTGGTGGCGGCGGCACTCACTCTGGCAGCCCGCTCC 413  
  
Db 282 CCCACGCCAGCCATGGTTGTCCCGCCATACGCAGTAGTTGAGTGTGGTCGGAAGGCCAG 341  
|||||  
Cp 412 CCCACGCCAGCCATGGTTGTCCCGCCATACGCAGTAGTTGAGTGTGGTCGGAAGGCCAG 353  
  
Db 342 CCAGGCCAGGTAGGGGTAGAGCAGCGGGCGGCCAGCGGGCTCACCTGGTACCAGGCCAC 401  
|||||  
Cp 352 CCAGGCCAGGTAGGGGTAGAGCAGCGGGCGGCCAGCGGGCTCACCTGGTACCAGGCCAC 293  
  
Db 402 GGAGTTGNCCTCCCGCCGCTGACCCAGCAGGAGATTACCAAGGCCAGCCCATTTG 461  
|||||  
Cp 292 GGAGTGGCTGCCCGCCGCTGACCCAGCAGGAGATCCACCAAGGCCAGCCCATTTG 233  
  
Db 462 TCGGGCACCAAAGAAGATGGGGGC-ATGCC-AGTTCAGGGCCAGCTGCC-AGTGTANAG 518  
|||||  
Cp 232 TCGGGCACCAAAGAAGATGGGGGCCATGCCAGTTTCAGGGCCAGCTGCCAGTGTAGAG 173  
  
Db 519 GCCAAGGG 526  
|||  
Cp 172 GCCCAGGG 165

RESULT 9  
LOCUS AA479680 541 bp mRNA EST 09-NOV-1997  
DEFINITION zu43a07.s1 Soares ovary tumor N6HOT Homo sapiens cDNA clone 740724  
3' similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE  
RECEPTOR (HUMAN);, mRNA sequence.  
ACCESSION AA479680  
NID g2205566  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 541)  
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,  
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,  
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,  
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.  
TITLE WashU-NCI human EST Project  
JOURNAL Unpublished (1997)  
COMMENT





Cp 472 ACCTGCTGGTGCAGCTGCAGTCCCTGGTGGCCGGCCGCACTCACTCTGGCAGCCCGCGTCC 413

Db 289 CCCAGCCAGCCATGTTGTCCCGCCATACGCAGTAGTGTAGTGTGGTGGCGAAGGCCAG 348  
|||||

Cp 412 CCCAGCCAGCCATGTTGTCCCGCCATACGCAGTAGTGTAGTGTGGTGGTGAAGGCCAG 353  
|||||

Db 349 CCAGCCAGGTAGGGGTAGAGAGCGGGCGGCCAGCGGGCTGCACCTGGTAGAGGCCA 408  
|||||

Cp 352 CCAGCCAGGTAGGGGTAGAGAGCGGGCGGCCAGCGGGCT-CACCTGGTAGCAGGCCA 294  
|||||

Db 409 CGTAGGTGNNCCGGCCGCGCCACTGACCCAGGAGATTACCAAGGCCCGCCAGCCATT 468  
|||||

Cp 293 CGTAGTGGCTGCCGCGCGCCACTGACCCAGGAGATCCACCAAGGCCCGCCAGCCATT 234  
|||||

Db 469 GTTCGGNAGCAAGAAGATGGGGGGCATGAC-ATTTAAGGGCAAGTTGCCCAAGTGTAGA 527  
|| ||

Cp 233 GTCGGCACCAAGAAGATGGGGGGCCATGCCCAAGTTCAGGGCCAGCTGCCCAAGTGTAGA 174  
|||||

Db 528 GGGCCAGGG 536  
|| |||||

Cp 173 GGCCAGGG 165

RESULT 11

LOCUS AA456315 506 bp mRNA EST 06-JUN-1997

DEFINITION aal3g10.s1 Soares NhHMPu S1 Homo sapiens cDNA clone 813186 3', similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR (HUMAN);, mRNA sequence.

ACCESSION AA456315

NID 92179525

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 506)

AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.

TITLE WashU-Merck EST Project 1997

JOURNAL Unpublished (1997)

COMMENT Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -41ml3 fwd. Et from Amersham  
High quality sequence stop: 443.  
Location/Qualifiers  
1..506  
/organism="Homo sapiens"  
/note="Organ: mixed (see below); Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbHM, pregnant uterus NbHPU, and fetal heart NbHHL9W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."  
/db\_xref="taxon:9606"  
/clone="813186"  
/clone\_lib="Soares NhHMPu S1"  
/tissue\_type="Pooled human melanocyte, fetal heart, and

pregnant uterus"  
/lab\_host="DH10B"  
complement(<1..>506)  
/db\_xref="GDB:5044096"

BASE COUNT 114 a 150 c 158 g 83 t 1 others

ORIGIN

Query Match 65.0%; Score 424; DB 7; Length 506;  
Best Local Similarity 96.9%; Pred. No. 0.00e+00;  
Matches 463; Conservative 0; Mismatches 10; Indels 5; Gaps 5;

Db 34 AAATTCCATGTTCCAAGAACAATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCG 93  
|||||

Cp 652 AAATTCCATGTTCCAAGAACAATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCG 593  
|||||

Db 94 GTGGGGCTCAGGTGGGGCAACCTCTGAAGCTCTGTGACCCCTGGGCCAAGGCCCTGAC 153  
|||||

Cp 592 GTGGGGCTCAGGTGGGGCACCTCTGAAGCTCTGTGACCCCTGGGCCAAGGCCCTGAC 533  
|||||

Db 154 AGACTAGCAGGCCCGCAGTGGTCATGAAGCGTGACGGCCACCATCATCACAAGCGTGATGGC 213  
|||||

Cp 532 AGACTAGCAGGCCCGCAGTGGTCATGAAGCGTGACGGCCACCATCATCACAAGCGTGATGGC 473  
|||||

Db 214 ACCTGCTGGTGAGGCTGCAGTCCCTGGTGGCCGGGCACACTCACTCTGGCAGCCCGGTCC 273  
|||||

Cp 472 ACCTGCTGGTGAGGCTGCAGTCCCTGGTGGCCGGGCACACTCACTCTGGCAGCCCGGTCC 413  
|||||

Db 274 CCATGCCAGCCATGTTGTCCGCATACGCAGTAGTGTAGTGTGGTGCAGAGGCCAG 333  
|||||

Cp 412 CCACGCCAGCCATGTTGTCCGCATACGCAGTAGTGTAGTGTGGTGCAGAGGCCAG 353  
|||||

Db 334 CCAGGCCAGGTAGGGGTAGAGCAGGGGGCGGCCAGCGGGCTCACCTGGTACCAGGCCAC 393  
|||||

Cp 352 CCAGGCCAGGTAGGGGTAGAGCAGGGGGCGGCCAGCGGGCTCACCTGGTACCAGGCCAC 293  
|||||

Db 394 GGTACTGGGN-CCGCCGCCCTCACTGACCAGCAGGAGATCCACCAAGGCCAGCC-ATTG 451  
|||||

Cp 292 GGTAGTGGCTGCCGCCGCCCTCACTGACCAGCAGGAGATCCACCAAGGCCAGCCATTG 233  
|||||

Db 452 TCGGGCAC-AAAGAAGATGGGGGGCCATGCC-AGTTTCAGAGGCAGCTGCC-AGTGTAG 506  
|||||

Cp 232 TCGGGCACCAAGAAGATGGGGGGCCATGCCAGTTTCAGGGCCAGCTGCCAGTGTAG 175  
|||||

RESULT 12

LOCUS AA036726 551 bp mRNA EST 10-MAY-1997

DEFINITION zk30b07.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 472021 3' similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR (HUMAN);, mRNA sequence.

ACCESSION AA036726

NID g1509983

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 551)

AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.

TITLE WashU-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 695 Std Error: 0.00  
Seq primer: -40ml3 fwd. from Amersham  
High quality sequence stop: 321.

FEATURES  
source

1. .551  
/organism="Homo sapiens"  
/note="Organ: uterus; Vector: pT7T3-Pac; Site\_1: Not I;  
Site\_2: Eco RI; 1st strand cDNA was primed with a Not I -  
oligo(dT) primer [5',  
AACTGGAGAAATTCGGCGCCGCTTTTCTTTTCTTTT 3']  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT7T3 vector. Library  
went through one round of normalization. Library  
constructed by M. Fatima Bonaldo."  
/db\_xref="taxon:9606"  
/clone="472021"  
/clone\_lib="Soares pregnant uterus NbHPU"  
/sex="female"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
complement(<1. .>551)  
/db\_xref="GDB:3758023"

mRNA

BASE COUNT 123 a 160 c 164 g 96 t 8 others  
ORIGIN

Query Match 63.5%; Score 414; DB 5; Length 551;  
Best Local Similarity 93.5%; Pred. No. 0.00e+00;  
Matches 491; Conservative 0; Mismatches 24; Indels 10; Gaps 10;

Db 30 AAATTCCATGTTCCAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 89  
|||||  
Cp 652 AAATTCCATGTTCCAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 593  
  
Db 90 GTGGGGCTCAGGTGGGGCCACTCTGAAGCTCTGTGACCCCTGGGCCAAGGCTG-GAC 148  
|||||  
Cp 592 GTGGGGCTCAGGTGGGGCCACTCTGAAGCTCTGTGACCCCTGGGCCAAGGCCCTGAC 533  
  
Db 149 AGACTAGCAGGCCCGCAGTGGTCATGAAAGCGTGACGGCCACACATCACAAAGCGTGATGGC 208  
|||||  
Cp 532 AGACTAGCAGGCCCGCAGTGGTCATGAAAGCGTGACGGCCACACATCACAAAGCGTGATGGC 473  
  
Db 209 ACCTGCTGGT-CAGCTGCAGTCCCTGTTGGGCGGGGCACCTCACTCTGGCAGCCGC-GTCC 266  
|||||  
Cp 472 ACCTGCTGGTGACGCTGCAGTCCCTGTTGGGCGGGGCACCTCACTCTGGCAGCCGCTCC 413  
  
Db 267 CCCACGCCAGCCATGGTTGTCCGCCATACGCAGTAGTTAGTGTGGTCTGTAAGGCCAG 326  
|||||  
Cp 412 CCCACGCCAGCCATGGTTGTCCGCCATACGCAGTAGTTAGTGTGGTCTGTAAGGCCAG 353  
  
Db 327 CCAGCGANGGTAAGGGGTAGAGAGCNGGGCCAGCGGCTCACCTGGTACCAAGGCCA 386  
|||||  
Cp 352 CCAGGCCAGGTA-GGGGTAGAGAGCGGGCGGGCCAGCGGCTCACCTGGTACCAAGGCCA 294  
  
Db 387 CGGTACTGTGANCAGCCGNCNCCACTTGACAGCAGGAGATCCACCAAGGCCAGNCCATT 446  
|||||  
Cp 293 CGGTAGTGGCTGCCGCGCCGCCACT-GACCAGCAGGAGATCCACCAAGGCCAGCCCAT 235  
  
Db 447 TTGTTCCGGCACCAAGAAGATGGCGGCCATGCCAGTTTCAGGGCCCAAGCTGCCATGTN 506  
|||||  
Cp 234 T-GT-CGGGCACCAAGAAGATGGGGGCCATGCCCAGTTCAGGGCCA-GCTGCCCACT- 179  
  
Db 507 GTAGAGCCCCAGGGGAACACAGNCTTCTTTGTTGANGCCTTCCA 551  
|||||  
Cp 178 GTAGAGCCCCAGGGGAACACAGCCTTCTCTGT-GAAGCCTCCCCA 135

RESULT 13  
LOCUS AA838533 505 bp mRNA EST 18-MAR-1998  
DEFINITION oe39all.s1 NCI\_CGAP\_Pr25 Homo sapiens cDNA clone IMAGE:1410908  
similar to gb:U12421\_cds1 PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR  
(HUMAN); mRNA sequence.

ACCESSION  
NID  
KEYWORDS  
SOURCE  
ORGANISM

AA838533  
g2913332  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 505)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert\_Strausberg@nih.gov  
Tissue Procurement: Suzanne L. Topalian, M.D., Robert K. Bright,  
Ph.D.  
CDNA Library Preparation: Stratagene, Inc.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 894 Std Error: 0.00  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 323.

FEATURES  
source

1. .505  
/organism="Homo sapiens"  
/note="Organ: prostate; Vector: Bluescript SK-; Site\_1:  
EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer:  
Oligo dt. Normal prostate epithelial cell line (HPV  
immortalized). 5' adaptor sequence: 5' GAATTCGGCAGG 3'  
3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'  
Average insert size: 1.1 kb."  
/db\_xref="taxon:9606"  
/clone="IMAGE:1410908"  
/clone\_lib="NCI\_CGAP\_Pr25"  
/tissue\_type="epithelium (cell line)"  
/lab\_host="SOLR (kanamycin resistant)"

BASE COUNT 112 a 148 c 156 g 89 t  
ORIGIN

Query Match 62.1%; Score 405; DB 16; Length 505;  
Best Local Similarity 94.9%; Pred. No. 0.00e+00;  
Matches 444; Conservative 0; Mismatches 21; Indels 3; Gaps 3;

Db 41 AAATTCCATGTTCCAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 100  
|||||  
Cp 652 AAATTCCATGTTCCAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCGG 593  
  
Db 101 GTGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCTTGAC 160  
|||||  
Cp 592 GTGGGGCTCAGGTGGGGCCACCTCTGAAGCTCTGCTGACCCCTGGGCCAAGGCTTGAC 533  
  
Db 161 AGACTAGCAGGCCCGCAGTGGTCATGAAAGCGTGACGGCCACCATCACAAAGCGTGATGGC 220  
|||||  
Cp 532 AGACTAGCAGGCCCGCAGTGGTCATGAAAGCGTGACGGCCACCATCACAAAGCGTGATGGC 473  
  
Db 221 ACCTGCTGGTCGAGTCGAGTCCCTGTTGGTGGCGGGGCACCTCACTCTGGCAGCCCGTCC 280  
|||||  
Cp 472 ACCTGCTGGTCGAGTCGAGTCCCTGTTGGTGGCGGGGCACCTCACTCTGGCAGCCCGTCC 413  
  
Db 281 CCCACGCCAGCCATGTTGTCCGCCATACGCAGTGTGAGTGTGTCGGAAG-CCAG 339  
|||||  
Cp 412 CCCACGCCAGCCATGTTGTCCGCCATACGCAGTGTGAGTGTGTCGGAAGGCCAG 353  
  
Db 340 CCAGGCAAGGTAGGGGTAGAGCAGCGGGCGGC-AGCGGGCTCACCTGGTACCAAGGCCAC 398  
|||||  
Cp 352 CCAGGCCAGGTAGGGGTAGAGCAGCGGGCGGCACCGGGCTCACCTGGTACCAAGGCCAC 293







/clone="IMAGE:1324464"  
/clone\_lib="NCI\_CGAP\_Kid5"  
/tissue\_type="2 pooled tumors (clear cell type)"  
/lab\_host="DH10B"

BASE COUNT 111 a 144 c 151 g 101 t  
ORIGIN

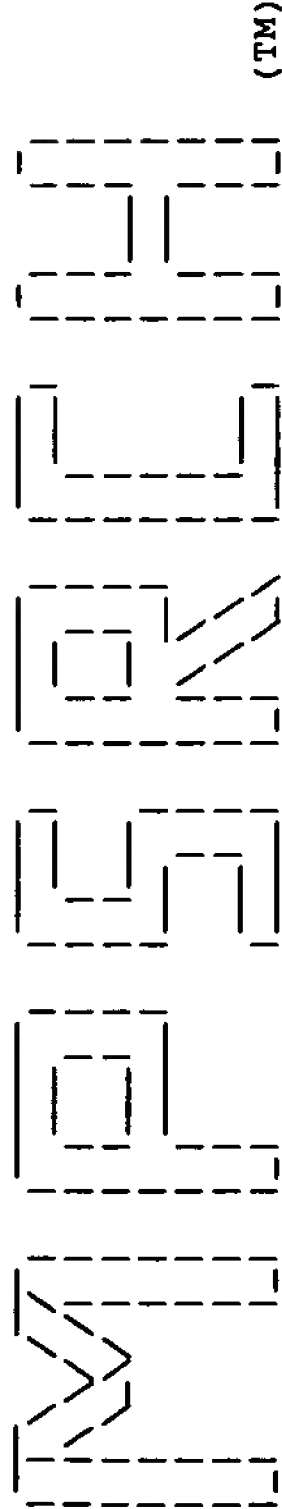
Query Match 60.4%; Score 394; DB 13; Length 507;  
Best Local Similarity 96.5%; Pred. No. 0.00e+00;  
Matches 435; Conservative 0; Mismatches 11; Indels 5; Gaps 5;

Db	53	AAATTCCATGTTCCCAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCCG	112
Cp	652	AAATTCCATGTTCCCAAGAACATGCTCTAAGCATGCAGAAAGCACAGGACACTGCTCCCCG	593
Db	113	GTGGGGGCTCAGGTGGGCCACCTCTGAAGCTCTGTGACCCCTGGGCCAAGCCTGTAC	172
Cp	592	GTGGGGGCTCAGGTGGGCCACCTCTGAAGCTCTGTGACCCCTGGGCCAAGCCTGTAC	533
Db	173	AGACTAGCAGGCCCATGGTGCATGAAAGCGTGACGGCCACCATCACAAAGCGTGATGGC	232
Cp	532	AGACTAGCAGGCCCATGGTGCATGAAAGCGTGACGGCCACCATCACAAAGCGTGATGGC	473
Db	233	ACCTGCTGGTGAGTCTGCAGTCCCTGGTGGGCCGGSCACTCACTCTGGCAGCCGCCGTCC	292
Cp	472	ACCTGCTGGTGAGTCTGCAGTCCCTGGTGGGCCGGSCACTCACTCTGGCAGCCGCCGTCC	413
Db	293	CCCACGCCAGCCATGGTTGTCCCGCCCATACGCAGTAGTTGAGTGTGGTCGTGAAGGCCAG	352
Cp	412	CCCACGCCAGCCATGGTTGTCCCGCCCATACGCAGTAGTTGAGTGTGGTCGTGAAGGCCAG	353
Db	353	CCAGGCCAGGTAGGGGTAGAGCAGGGCGGGCCAGCGGGCTACACCTGGTACAGGCCCA	412
Cp	352	CCAGGCCAGGTAGGGGTAGAGCAGGGCGGGCCAGCGGGCT-CACCTGGTACAGGCCCA	294
Db	413	CGGTAGGTGACAC-GCCGCCCCACTGACCAGCAGGAGATCCAC-AAGGCC-AGCCCATTT	469
Cp	293	CGGTAGTGGCTGCCCGCCGCCCACTGACCAGCAGGAGATCCACCAAGGCCCCAGCCCATTT	234
Db	470	GTCGGGCACCAA-GAAGATGGGGGCCCATGC	499
Cp	233	GTCGGGCACCAAAGATGGGGGGCCATGC	203

Search completed: Tue Jul 21 16:04:55 1998  
Job time : 776 secs.

**THIS PAGE BLANK (USPTO)**

\*\*\*\*\*



\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Jul 21 14:52:56 1998; MasPar time 7.43 Seconds  
Tabular output not generated. 349.017 Million cell updates/sec

Title: >US-09-047-652A-3  
Description: (1-169) from US09047652A.pep  
Perfect Score: 1170  
Sequence: 1 XXXXXXXXXXXXXXXXXX.....LNYCVWRDNGHWRGRRRLPE 169

Scoring table: PAM 150  
Gap 11

Searched: 124785 seqs, 15338987 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq31-2  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27

Statistics: Mean 31.057; Variance 140.986; scale 0.220

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	100	8.5	483	20	W12136 Alpha-amylase variant	5.89e+00
2	100	8.5	483	20	W12143 Alpha-amylase variant	5.89e+00
3	100	8.5	483	20	W12137 Alpha-amylase variant	5.89e+00
4	100	8.5	483	20	W12138 Alpha-amylase variant	5.89e+00
5	100	8.5	483	20	W12139 Alpha-amylase variant	5.89e+00
6	100	8.5	483	20	W12140 Alpha-amylase variant	5.89e+00
7	100	8.5	483	20	W12115 Alpha-amylase variant	5.89e+00
8	100	8.5	483	20	W12116 Alpha-amylase variant	5.89e+00
9	100	8.5	483	20	W12134 Alpha-amylase variant	5.89e+00
10	100	8.5	483	20	W12133 Alpha-amylase variant	5.89e+00
11	100	8.5	485	27	W31500 Bacillus sp. alpha am	5.89e+00
12	100	8.5	485	27	W31499 Bacillus sp. alpha am	5.89e+00
13	100	8.5	485	20	W12130 Alpha-amylase variant	5.89e+00
14	100	8.5	485	20	W12131 Alpha-amylase variant	5.89e+00
15	100	8.5	485	16	R81836 Bacillus sp. alkaline	5.89e+00
16	100	8.5	485	20	W12127 Alpha-amylase variant	5.89e+00
17	100	8.5	485	20	W12125 Alpha-amylase variant	5.89e+00
18	100	8.5	485	20	W12144 Alpha-amylase variant	5.89e+00
19	100	8.5	485	20	W12110 Alpha-amylase variant	5.89e+00

20	100	8.5	485	20	W12109 Alpha-amylase variant	5.89e+00
21	100	8.5	485	20	W12108 Alpha-amylase variant	5.89e+00
22	100	8.5	485	20	W12104 Alpha-amylase variant	5.89e+00
23	100	8.5	485	20	W12101 Alpha-amylase variant	5.89e+00
24	100	8.5	485	20	W12111 Alpha-amylase variant	5.89e+00
25	100	8.5	485	20	W12117 Alpha-amylase variant	5.89e+00
26	100	8.5	485	16	R81835 Bacillus sp. alkaline	5.89e+00
27	100	8.5	485	20	W12112 Alpha-amylase variant	5.89e+00
28	100	8.5	485	20	W12102 Alpha-amylase variant	5.89e+00
29	100	8.5	485	20	W12126 Alpha-amylase variant	5.89e+00
30	100	8.5	485	20	W12120 Alpha-amylase variant	5.89e+00
31	100	8.5	485	20	W12098 Alpha-amylase variant	5.89e+00
32	100	8.5	485	20	W12113 Alpha-amylase variant	5.89e+00
33	100	8.5	485	20	W12114 Alpha-amylase variant	5.89e+00
34	100	8.5	485	20	W12128 Alpha-amylase variant	5.89e+00
35	100	8.5	485	20	W12129 Alpha-amylase variant	5.89e+00
36	100	8.5	485	20	W12100 Alpha-amylase variant	5.89e+00
37	100	8.5	485	20	W12099 Alpha-amylase variant	5.89e+00
38	100	8.5	485	20	W12105 Alpha-amylase variant	5.89e+00
39	100	8.5	485	20	W12106 Alpha-amylase variant	5.89e+00
40	100	8.5	485	20	W12103 Alpha-amylase variant	5.89e+00
41	100	8.5	485	20	W12107 Alpha-amylase variant	5.89e+00
42	100	8.5	485	20	W12119 Alpha-amylase variant	5.89e+00
43	100	8.5	485	20	W12132 Alpha-amylase variant	5.89e+00
44	100	8.5	485	20	W12122 Alpha-amylase variant	5.89e+00
45	100	8.5	516	24	W11326 Alkaline liquefying a	5.89e+00

ALIGNMENTS

RESULT 1  
ID W12136 standard; protein; 483 AA.  
AC W12136;  
DT 08-APR-1997 (first entry)  
DE Alpha-amylase variant L351C+M430C+Y243F+delta183+delta184+Q391E+K444Q.  
KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
KW textile desizing; papermaking; beer-making; ethanol production;  
KW sweetener.  
OS Synthetic.  
FH Key  
FT misc\_difference 243 Location/Qualifiers  
FT misc\_difference 351 /label= Y243F  
FT misc\_difference 391 /label= L351C  
FT misc\_difference 430 /label= Q391E  
FT misc\_difference 444 /label= M430C  
FT misc\_difference 444 /label= K444Q  
PN WO9623873-A1.  
PD 08-AUG-1996.  
PF 05-FEB-1996; DK0056.  
PR 03-FEB-1995; DK-000126.  
PR 29-MAR-1995; DK-000336.  
PR 29-SEP-1995; DK-001097.  
PR 06-OCT-1995; DK-001121.  
PA (NOVO ) NOVO-NORDISK AS.  
PI Bisgard-frantzen H, Borchert T, Svendsen A;  
DR WPI: 96-371423/37  
PT Alpha-amylase variants - with improved thermal and oxidation  
PS stability and reduced calcium ion dependency  
PS Example 2; ; 11pp; English.  
CC W12098-W12144 represent alpha-amylase variants of the invention. The  
CC variants of the invention were created using site directed, or random,  
CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,  
CC W12142 and W12144 are specifically variants of the alkaphilic Bacillus  
CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have  
CC improved thermal stability (such as at temperatures in the range of 40-70  
CC degrees Celcius), and/or oxidation stability, and/or reduced calcium ion  
CC dependency. The variants can also have increased alpha-amylolytic

CC activity (especially at pH values in the range of 8.5-10.5), and improved  
CC binding of a particular substrate. These variant alpha-amylases also  
CC possess improved specificity to a particular substrate, and/or improved  
CC specificity with respect to cleavage of substrate. These sequences can  
CC be used in detergent and washing compositions, and for textile desizing.  
CC The alpha-amylase variants can also be used in papermaking and  
CC beer-making processes. These variants can also be used in the production  
CC of sweeteners and ethanol from starch.  
SQ Sequence 483 AA;

Query Match 8.5%; Score 100; DB 20; Length 483;  
Best Local Similarity 40.0%; Pred. No. 5.89e+00;  
Matches 14; Conservative 8; Mismatches 10; Indels 3; Gaps 3;

Db 43 w1-ppawkgtsqndvgvgaydly-dlgefngkgtv 75  
|:| | | | : |||:| : |||:| |  
QY 47 WVLGPVW-GTLYSAMGYGSYLVWKELGGFTEKAVV 80

RESULT 2  
ID W12143 standard; protein; 483 AA.  
AC W12143;

DT 08-APR-1997 (first entry)  
DE Alpha-amylase variant delta182 + delta184.  
KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
KW textile desizing; papermaking; beer-making; ethanol production;  
KW sweetener.  
OS Synthetic.  
PN WO9623873-A1.  
PD 08-AUG-1996.  
PF 05-FEB-1996; DK0056.  
PR 03-FEB-1995; DK-000126.  
PR 29-MAR-1995; DK-000336.  
PR 29-SEP-1995; DK-001097.  
PR 06-OCT-1995; DK-001121.  
PA (NOVO ) NOVO-NORDISK AS.  
PI Bisgard-frantzen H, Borchert T, Svendsen A;  
DR WPI; 96-371423/37.  
PT Alpha-amylase variants - with improved thermal and oxidation  
PT stability and reduced calcium ion dependency  
PS Example 3; ; 11lpp; English.

CC W12098-W12144 represent alpha-amylase variants of the invention. The  
CC variants of the invention were created using site directed, or random,  
CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
CC represented by W12955, W12956, R81835 and R81836. W12137-W12140, and  
CC W12143 are specifically variants of the alkaphilic Bacillus strain NCIB  
CC 12513 alpha-amylase shown in R81836. These variants can have improved  
CC thermal stability (such as at temperatures in the range of 40-70 degrees  
CC Celcius), and/or oxidation stability, and/or reduced calcium ion  
CC dependency. The variants can also have increased alpha-amylolytic  
CC activity (especially at pH values in the range of 8.5-10.5), and improved  
CC binding of a particular substrate. These variant alpha-amylases also  
CC possess improved specificity to a particular substrate, and/or improved  
CC specificity with respect to cleavage of substrate. These sequences can  
CC be used in detergent and washing compositions, and for textile desizing.  
CC The alpha-amylase variants can also be used in papermaking and  
CC beer-making processes. These variants can also be used in the production  
CC of sweeteners and ethanol from starch.  
SQ Sequence 483 AA;

Query Match 8.5%; Score 100; DB 20; Length 483;  
Best Local Similarity 40.0%; Pred. No. 5.89e+00;  
Matches 14; Conservative 8; Mismatches 10; Indels 3; Gaps 3;

Db 43 w1-ppawkgtsqndvgvgaydly-dlgefngkgtv 75  
|:| | | | : |||:| : |||:| |  
QY 47 WVLGPVW-GTLYSAMGYGSYLVWKELGGFTEKAVV 80

RESULT 3  
ID W12137 standard; protein; 483 AA.  
AC W12137;

DT 08-APR-1997 (first entry)  
DE Alpha-amylase variant delta183 + delta184.  
KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
KW textile desizing; papermaking; beer-making; ethanol production;  
KW sweetener.  
OS Synthetic.  
PN WO9623873-A1.  
PD 08-AUG-1996.  
PF 05-FEB-1996; DK0056.  
PR 03-FEB-1995; DK-000126.  
PR 29-MAR-1995; DK-000336.  
PR 29-SEP-1995; DK-001097.  
PR 06-OCT-1995; DK-001121.  
PA (NOVO ) NOVO-NORDISK AS.  
PI Bisgard-frantzen H, Borchert T, Svendsen A;  
DR WPI; 96-371423/37.  
PT Alpha-amylase variants - with improved thermal and oxidation  
PT stability and reduced calcium ion dependency  
PS Example 2; ; 11lpp; English.

CC W12098-W12144 represent alpha-amylase variants of the invention. The  
CC variants of the invention were created using site directed, or random,  
CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
CC represented by W12955, W12956, R81835 and R81836. W12137-W12140, and  
CC W12143 are specifically variants of the alkaphilic Bacillus strain NCIB  
CC 12513 alpha-amylase shown in R81836. These variants can have improved  
CC thermal stability (such as at temperatures in the range of 40-70 degrees  
CC Celcius), and/or oxidation stability, and/or reduced calcium ion  
CC dependency. The variants can also have increased alpha-amylolytic  
CC activity (especially at pH values in the range of 8.5-10.5), and improved  
CC binding of a particular substrate. These variant alpha-amylases also  
CC possess improved specificity to a particular substrate, and/or improved  
CC specificity with respect to cleavage of substrate. These sequences can  
CC be used in detergent and washing compositions, and for textile desizing.  
CC The alpha-amylase variants can also be used in papermaking and  
CC beer-making processes. These variants can also be used in the production  
CC of sweeteners and ethanol from starch.  
SQ Sequence 483 AA;

Query Match 8.5%; Score 100; DB 20; Length 483;  
Best Local Similarity 40.0%; Pred. No. 5.89e+00;  
Matches 14; Conservative 8; Mismatches 10; Indels 3; Gaps 3;

Db 43 w1-ppawkgtsqndvgvgaydly-dlgefngkgtv 75  
|:| | | | : |||:| : |||:| |  
QY 47 WVLGPVW-GTLYSAMGYGSYLVWKELGGFTEKAVV 80

RESULT 4  
ID W12138 standard; protein; 483 AA.  
AC W12138;

DT 08-APR-1997 (first entry)  
DE Alpha-amylase variant delta181 + delta183.  
KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
KW textile desizing; papermaking; beer-making; ethanol production;  
KW sweetener.  
OS Synthetic.  
PN WO9623873-A1.  
PD 08-AUG-1996.  
PF 05-FEB-1996; DK0056.  
PR 03-FEB-1995; DK-000126.  
PR 29-MAR-1995; DK-000336.  
PR 29-SEP-1995; DK-001097.  
PR 06-OCT-1995; DK-001121.  
PA (NOVO ) NOVO-NORDISK AS.  
PI Bisgard-frantzen H, Borchert T, Svendsen A;  
DR WPI; 96-371423/37.  
PT Alpha-amylase variants - with improved thermal and oxidation  
PT stability and reduced calcium ion dependency  
PS Example 2; ; 11lpp; English.

CC W12098-W12144 represent alpha-amylase variants of the invention. The  
CC variants of the invention were created using site directed, or random,  
CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
CC represented by W12955, W12956, R81835 and R81836. W12137-W12140, and  
CC W12143 are specifically variants of the alkaphilic Bacillus strain NCIB  
CC 12513 alpha-amylase shown in R81836. These variants can have improved  
CC thermal stability (such as at temperatures in the range of 40-70 degrees  
CC Celcius), and/or oxidation stability, and/or reduced calcium ion  
CC dependency. The variants can also have increased alpha-amylolytic  
CC activity (especially at pH values in the range of 8.5-10.5), and improved  
CC binding of a particular substrate. These variant alpha-amylases also  
CC possess improved specificity to a particular substrate, and/or improved  
CC specificity with respect to cleavage of substrate. These sequences can  
CC be used in detergent and washing compositions, and for textile desizing.  
CC The alpha-amylase variants can also be used in papermaking and  
CC beer-making processes. These variants can also be used in the production  
CC of sweeteners and ethanol from starch.  
SQ Sequence 483 AA;

Query Match 8.5%; Score 100; DB 20; Length 483;  
Best Local Similarity 40.0%; Pred. No. 5.89e+00;  
Matches 14; Conservative 8; Mismatches 10; Indels 3; Gaps 3;

Db 43 w1-ppawkgtsqndvgvgaydly-dlgefngkgtv 75  
|:| | | | : |||:| : |||:| |  
QY 47 WVLGPVW-GTLYSAMGYGSYLVWKELGGFTEKAVV 80

RESULT 3  
ID W12137 standard; protein; 483 AA.  
AC W12137;



CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
CC represented by W12955, W12956, R81835 and R81836. W12137-W12140, and  
CC W12143 are specifically variants of the alkaphilic Bacillus strain NCIB  
CC 12513 alpha-amylase shown in R81836. These variants can have improved  
CC thermal stability (such as at temperatures in the range of 40-70 degrees  
CC Celcius), and/or oxidation stability, and/or reduced calcium ion  
CC dependency. The variants can also have increased alpha-amylolytic  
CC activity (especially at pH values in the range of 8.5-10.5), and improved  
CC binding of a particular substrate. These variant alpha-amylases also  
CC possess improved specificity to a particular substrate, and/or improved  
CC specificity with respect to cleavage of substrate. These sequences can  
CC be used in detergent and washing compositions, and for textile desizing.  
CC The alpha-amylase variants can also be used in papermaking and  
CC beer-making processes. These variants can also be used in the production  
CC of sweeteners and ethanol from starch.  
SQ Sequence 483 AA;

Query Match 8.5%; Score 100; DB 20; Length 483;  
Best Local Similarity 40.0%; Pred. No. 5.89e+00;  
Matches 14; Conservative 8; Mismatches 10; Indels 3; Gaps 3;

Db 43 wi-ppawkgtsqndvgygaydly-dlgefngkgtv 75  
QY 47 WVLGPVW-GTLYSAMYGSYLVWVKELGGFTEKAVV 80

RESULT 5  
ID W12139 standard; protein; 483 AA.  
AC W12139;  
DT 08-APR-1997 (first entry)  
DE Alpha-amylase variant delta181 + delta182.  
KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
KW textile desizing; papermaking; beer-making; ethanol production;  
KW sweetener.  
OS Synthetic.  
PN WO9623873-A1.  
PD 08-AUG-1996.  
PF 05-FEB-1996; DK0056.  
PR 03-FEB-1995; DK-000126.  
PR 29-MAR-1995; DK-000336.  
PR 29-SEP-1995; DK-001097.  
PR 06-OCT-1995; DK-001121.  
PA (NOVO ) NOVO-NORDISK AS.  
PI Bisgard-frantzen H, Borchert T, Svendsen A;  
DR WPI; 96-371423/37.  
PT Alpha-amylase variants - with improved thermal and oxidation  
PT stability and reduced calcium ion dependency  
PS Example 2; ; 11lpp; English.  
CC W12098-W12144 represent alpha-amylase variants of the invention. The  
CC variants of the invention were created using site directed, or random,  
CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
CC represented by W12955, W12956, R81835 and R81836. W12137-W12140, and  
CC W12143 are specifically variants of the alkaphilic Bacillus strain NCIB  
CC 12513 alpha-amylase shown in R81836. These variants can have improved  
CC thermal stability (such as at temperatures in the range of 40-70 degrees  
CC Celcius), and/or oxidation stability, and/or reduced calcium ion  
CC dependency. The variants can also have increased alpha-amylolytic  
CC activity (especially at pH values in the range of 8.5-10.5), and improved  
CC binding of a particular substrate. These variant alpha-amylases also  
CC possess improved specificity to a particular substrate, and/or improved  
CC specificity with respect to cleavage of substrate. These sequences can  
CC be used in detergent and washing compositions, and for textile desizing.  
CC The alpha-amylase variants can also be used in papermaking and  
CC beer-making processes. These variants can also be used in the production  
CC of sweeteners and ethanol from starch.  
SQ Sequence 483 AA;

Query Match 8.5%; Score 100; DB 20; Length 483;  
Best Local Similarity 40.0%; Pred. No. 5.89e+00;  
Matches 14; Conservative 8; Mismatches 10; Indels 3; Gaps 3;  
Db 43 wi-ppawkgtsqndvgygaydly-dlgefngkgtv 75

QY 47 WVLGPVW-GTLYSAMYGSYLVWVKELGGFTEKAVV 80  
RESULT 6  
ID W12140 standard; protein; 483 AA.  
AC W12140;  
DT 08-APR-1997 (first entry)  
DE Alpha-amylase variant delta183 + delta184 + M202L/I.  
KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
KW textile desizing; papermaking; beer-making; ethanol production;  
KW sweetener.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT misc\_difference 202  
FT /label= M202L/I  
PN WO9623873-A1.  
PD 08-AUG-1996.  
PF 05-FEB-1996; DK0056.  
PR 03-FEB-1995; DK-000126.  
PR 29-MAR-1995; DK-000336.  
PR 29-SEP-1995; DK-001097.  
PR 06-OCT-1995; DK-001121.  
PA (NOVO ) NOVO-NORDISK AS.  
PI Bisgard-frantzen H, Borchert T, Svendsen A;  
DR WPI; 96-371423/37.  
PT Alpha-amylase variants - with improved thermal and oxidation  
PT stability and reduced calcium ion dependency  
PS Example 2; ; 11lpp; English.  
CC W12098-W12144 represent alpha-amylase variants of the invention. The  
CC variants of the invention were created using site directed, or random,  
CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
CC represented by W12955, W12956, R81835 and R81836. W12137-W12140, and  
CC W12143 are specifically variants of the alkaphilic Bacillus strain NCIB  
CC 12513 alpha-amylase shown in R81836. These variants can have improved  
CC thermal stability (such as at temperatures in the range of 40-70 degrees  
CC Celcius), and/or oxidation stability, and/or reduced calcium ion  
CC dependency. The variants can also have increased alpha-amylolytic  
CC activity (especially at pH values in the range of 8.5-10.5), and improved  
CC binding of a particular substrate. These variant alpha-amylases also  
CC possess improved specificity to a particular substrate, and/or improved  
CC specificity with respect to cleavage of substrate. These sequences can  
CC be used in detergent and washing compositions, and for textile desizing.  
CC The alpha-amylase variants can also be used in papermaking and  
CC beer-making processes. These variants can also be used in the production  
CC of sweeteners and ethanol from starch.  
SQ Sequence 483 AA;

Query Match 8.5%; Score 100; DB 20; Length 483;  
Best Local Similarity 40.0%; Pred. No. 5.89e+00;  
Matches 14; Conservative 8; Mismatches 10; Indels 3; Gaps 3;  
Db 43 wi-ppawkgtsqndvgygaydly-dlgefngkgtv 75  
QY 47 WVLGPVW-GTLYSAMYGSYLVWVKELGGFTEKAVV 80  
RESULT 7  
ID W12115 standard; protein; 483 AA.  
AC W12115;  
DT 08-APR-1997 (first entry)  
DE Alpha-amylase variant delta181 + delta182.  
KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
KW textile desizing; papermaking; beer-making; ethanol production;  
KW sweetener.  
OS Synthetic.  
PN WO9623873-A1.  
PD 08-AUG-1996.  
PF 05-FEB-1996; DK0056.  
PR 03-FEB-1995; DK-000126.  
PR 29-MAR-1995; DK-000336.



DE 08-APR-1997 (first entry)  
DE Alpha-amylase variant R124P + delta183 + delta184.  
KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
KW textile desizing; papermaking; beer-making; ethanol production;  
KW sweetener.  
OS Synthetic.  
FH Key  
FT misc\_difference 124 Location/Qualifiers  
FT /label= R124P  
PN WO9623873-A1.  
PD 08-AUG-1996.  
PF 05-FEB-1996; DK0056.  
PR 03-FEB-1995; DK-000126.  
PR 29-MAR-1995; DK-000336.  
PR 29-SEP-1995; DK-001097.  
PR 06-OCT-1995; DK-001121.  
PA (NOVO ) NOVO-NORDISK AS.  
PI Bisgard-frantzen H, Borchert T, Svendsen A;  
PI WPI; 96-371423/37.  
DR Alpha-amylase variants - with improved thermal and oxidation  
PT stability and reduced calcium ion dependency  
PT Example 2; ; 111pp; English.  
CC W12098-W12144 represent alpha-amylase variants of the invention. The  
CC variants of the invention were created using site directed, or random,  
CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,  
CC W12142 and W12144 are specifically variants of the alkaphilic Bacillus  
CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have  
CC improved thermal stability (such as at temperatures in the range of 40-70  
CC degrees Celcius), and/or oxidation stability, and/or reduced calcium ion  
CC dependency. The variants can also have increased alpha-amylolytic  
CC activity (especially at pH values in the range of 8.5-10.5), and improved  
CC binding of a particular substrate. These variant alpha-amylases also  
CC possess improved specificity to a particular substrate, and/or improved  
CC specificity with respect to cleavage of substrate. These sequences can  
CC be used in detergent and washing compositions, and for textile desizing.  
CC The alpha-amylase variants can also be used in papermaking and  
CC beer-making processes. These variants can also be used in the production  
CC of sweeteners and ethanol from starch.  
SQ Sequence 483 AA;

Query Match 8.5%; Score 100; DB 20; Length 483;  
Best Local Similarity 40.0%; Pred. No. 5.89e+00;  
Matches 14; Conservative 8; Mismatches 10; Indels 3; Gaps 3;

Db 43 w1-ppawkgtsqndvgygaydly-dlgefngkgtv 75  
|: | | | | : : | | | | : : | | | | : |  
QY 47 WVLGPVW-GTLYSAMGYGSYLVWKELGGFTEKAVV 80

RESULT 11  
ID W31500 standard; protein; 485 AA.  
AC W31500;  
DT 08-APR-1998 (first entry)  
DE Bacillus sp. alpha amylase.  
KW Alpha amylase; hard surface cleaning; dishwashing; laundry.  
OS Bacillus sp.  
PN WO9732961-A2.  
PD 12-SEP-1997.  
PF 04-MAR-1997; U03635.  
PR 07-MAR-1996; WO-U03276.  
PA (PROC ) PROCTER & GAMBLE CO.  
PI Baeck AC, Jones LA, Ohtani R, Pramod K, Rai S,  
PI Showell MS, Ward G;  
PI WPI; 97-457524/42.  
DR Detergent compositions for hard surface cleaning and laundry use -  
PT contains Bacillus derived alpha amylase with improved  
PT thermostability, reduced calcium ion dependency etc.  
PS Claim 1; Pages 88-89; 97pp; English.  
CC The present sequence is a Bacillus sp. alpha amylase with a  
CC specific activity at least 25% higher than that of Termamyl (RTM)  
CC at 25 to 55 degrees C at pH 8 to 10, measured by the Phadebas (RTM)

CC activity test. It is of use in hard surface cleaning, hand or  
CC machine dishwashing and laundry at a temperature of 10 to 25  
CC degrees C. Improved cleaning, stain removal and fabric care are  
CC obtained by using it at a concentration of 0.00018 to 0.06%  
SQ Sequence 485 AA;  
  
Query Match 8.5%; Score 100; DB 27; Length 485;  
Best Local Similarity 40.0%; Pred. No. 5.89e+00;  
Matches 14; Conservative 8; Mismatches 10; Indels 3; Gaps 3;

Db 43 w1-ppawkgtsqndvgygaydly-dlgefngkgtv 75  
|: | | | | : : | | | | : : | | | | : |  
QY 47 WVLGPVW-GTLYSAMGYGSYLVWKELGGFTEKAVV 80

RESULT 12  
ID W31499 standard; protein; 485 AA.  
AC W31499;  
DT 08-APR-1998 (first entry)  
DE Bacillus sp. alpha amylase.  
KW Alpha amylase; hard surface cleaning; dishwashing; laundry.  
OS Bacillus sp.  
PN WO9732961-A2.  
PD 12-SEP-1997.  
PF 04-MAR-1997; U03635.  
PR 07-MAR-1996; WO-U03276.  
PA (PROC ) PROCTER & GAMBLE CO.  
PI Baeck AC, Jones LA, Ohtani R, Pramod K, Rai S,  
PI Showell MS, Ward G;  
PI WPI; 97-457524/42.  
DR Detergent compositions for hard surface cleaning and laundry use -  
PT contains Bacillus derived alpha amylase with improved  
PT thermostability, reduced calcium ion dependency etc.  
PS Claim 1; Pages 86-87; 97pp; English.  
CC The present sequence is a Bacillus sp. alpha amylase with a  
CC specific activity at least 25% higher than that of Termamyl (RTM)  
CC at 25 to 55 degrees C at pH 8 to 10, measured by the Phadebas (RTM)  
CC activity test. It is of use in hard surface cleaning, hand or  
CC machine dishwashing and laundry at a temperature of 10 to 25  
CC degrees C. Improved cleaning, stain removal and fabric care are  
CC obtained by using it at a concentration of 0.00018 to 0.06%  
SQ Sequence 485 AA;

Query Match 8.5%; Score 100; DB 27; Length 485;  
Best Local Similarity 40.0%; Pred. No. 5.89e+00;  
Matches 14; Conservative 8; Mismatches 10; Indels 3; Gaps 3;

Db 43 w1-ppawkgtsqndvgygaydly-dlgefngkgtv 75  
|: | | | | : : | | | | : : | | | | : |  
QY 47 WVLGPVW-GTLYSAMGYGSYLVWKELGGFTEKAVV 80

RESULT 13  
ID W12130 standard; protein; 485 AA.  
AC W12130;  
DT 08-APR-1997 (first entry)  
DE Alpha-amylase variant E194Q.  
KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
KW textile desizing; papermaking; beer-making; ethanol production;  
KW sweetener.  
OS Synthetic.  
FH Key  
FT misc\_difference 194 Location/Qualifiers  
FT /label= E194Q  
PN WO9623873-A1.  
PD 08-AUG-1996.  
PF 05-FEB-1996; DK0056.  
PR 03-FEB-1995; DK-000126.  
PR 29-MAR-1995; DK-000336.  
PR 29-SEP-1995; DK-001097.  
PR 06-OCT-1995; DK-001121.  
PA (NOVO ) NOVO-NORDISK AS.



PI Bisgard-frantzen H, Borchert T, Svendsen A;  
DR WPI; 96-371423/37.  
PT Alpha-amylase variants - with improved thermal and oxidation  
PT stability and reduced calcium ion dependency  
PS Claim 11; ; 111pp; English.  
CC W12098-W12144 represent alpha-amylase variants of the invention. The  
CC variants of the invention were created using site directed, or random,  
CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,  
CC W12142 and W12144 are specifically variants of the alkaphilic Bacillus  
CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have  
CC improved thermal stability (such as at temperatures in the range of 40-70  
CC degrees Celcius), and/or oxidation stability, and/or reduced calcium ion  
CC dependency. The variants can also have increased alpha-amylolytic  
CC activity (especially at pH values in the range of 8.5-10.5), and improved  
CC binding of a particular substrate. These variant alpha-amylases also  
CC possess improved specificity to a particular substrate, and/or improved  
CC specificity with respect to cleavage of substrate. These sequences can  
CC be used in detergent and washing compositions, and for textile desizing.  
CC The alpha-amylase variants can also be used in papermaking and  
CC beer-making processes. These variants can also be used in the production  
CC of sweeteners and ethanol from starch.  
SQ Sequence 485 AA;

Query Match 8.5%; Score 100; DB 20; Length 485;  
Best Local Similarity 40.0%; Pred. No. 5.89e+00;  
Matches 14; Conservative 8; Mismatches 10; Indels 3; Gaps 3;

Db 43 wi-ppawkgtsqndvgygaydly-dlgefnqgkv 75  
I: | | | | : : | | | | : : | | | | : | | : | |  
QY 47 WVLGPVW-GTLYSAMGYGSYLVWKELGGFTEKAVV 80

RESULT 14  
ID W12131 standard; protein; 485 AA.  
AC W12131;  
DT 08-APR-1997 (first entry)  
DE Alpha-amylase variant N106D.  
KW Alpha-amylase; detergent; thermal stability; oxidation stability; enzyme;  
KW calcium ion dependency; alpha-amylolytic activity; washing composition;  
KW textile desizing; papermaking; beer-making; ethanol production;  
KW sweetener.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT misc\_difference 106  
FT /label= N106D  
PD W09623873-A1.  
PN 08-AUG-1996.  
PF 05-FEB-1996; DK0056.  
PR 03-FEB-1995; DK-000126.  
PR 29-MAR-1995; DK-000336.  
PR 29-SEP-1995; DK-001097.  
PR 06-OCT-1995; DK-001121.  
PA (NOVO ) NOVO-NORDISK AS.  
PI Bisgard-frantzen H, Borchert T, Svendsen A;  
DR WPI; 96-371423/37.  
PT Alpha-amylase variants - with improved thermal and oxidation  
PT stability and reduced calcium ion dependency  
PS Claim 11; ; 111pp; English.  
CC W12098-W12144 represent alpha-amylase variants of the invention. The  
CC mutagenesis of the DNA sequences encoding the parent alpha-amylases  
CC represented by W12955, W12956, R81835 and R81836. W12098-W12136, W12141,  
CC W12142 and W12144 are specifically variants of the alkaphilic Bacillus  
CC strain NCIB 12512 alpha-amylase shown in R81835. These variants can have  
CC improved thermal stability (such as at temperatures in the range of 40-70  
CC degrees Celcius), and/or oxidation stability, and/or reduced calcium ion  
CC dependency. The variants can also have increased alpha-amylolytic  
CC activity (especially at pH values in the range of 8.5-10.5), and improved  
CC binding of a particular substrate. These variant alpha-amylases also  
CC possess improved specificity to a particular substrate. These sequences can  
CC be used in detergent and washing compositions, and for textile desizing.

CC The alpha-amylase variants can also be used in papermaking and  
CC beer-making processes. These variants can also be used in the production  
CC of sweeteners and ethanol from starch.  
SQ Sequence 485 AA;

Query Match 8.5%; Score 100; DB 20; Length 485;  
Best Local Similarity 40.0%; Pred. No. 5.89e+00;  
Matches 14; Conservative 8; Mismatches 10; Indels 3; Gaps 3;

Db 43 wi-ppawkgtsqndvgygaydly-dlgefnqgkv 75  
I: | | | | : : | | | | : : | | | | : | | : | |  
QY 47 WVLGPVW-GTLYSAMGYGSYLVWKELGGFTEKAVV 80

RESULT 15  
ID R81836 standard; protein; 485 AA.  
AC R81836;  
DT 16-MAR-1996 (first entry)  
DE Bacillus sp. alkaline alpha-amylase.  
KW Alpha-amylase; enzyme; EC-3.2.1.1; detergent; surfactant; textile;  
KW beer; starch.  
OS Bacillus.  
PN W09526397-A1.  
PD 05-OCT-1995.  
PF 29-MAR-1995; DK0142.  
PR 29-MAR-1994; DK-000353.  
PR 03-NOV-1994; DK-001271.  
PR 03-FEB-1995; DK-000123.  
PA (NOVO ) NOVO-NORDISK AS.  
PI Bisgard-frantzen H, Ostergaard PR, Outtrup H, Rasmussen MD;  
PI Van DER ZEE P;  
DR WPI; 95-351318/45.  
DR N-PSDB; T00777.  
PT New alkaline Bacillus alpha-amylase - used in e.g. detergent  
PT compsns. starch liquefaction, textile desizing, starch modification  
PT or beer making  
PS Claim 5; Page 46-48; 65pp; English.  
CC This alpha-amylase protein is characterized by having a specific  
CC activity at least 25% higher than the specific activity of Termamyl  
CC at 25-55 deg C and a pH of 8-10. The enzyme can be used in  
CC detergent composition for starch liquefaction, the production of  
CC lignocellulosic materials, e.g. pulp, paper and cardboard from waste  
CC containing starch, for deinking recycled starch-coated, or starch-  
CC containing printed paper, to modify starch for papermaking, for  
CC textile desizing, and beer-making processes. This protein may be  
CC produced recombinantly.  
SQ Sequence 485 AA;

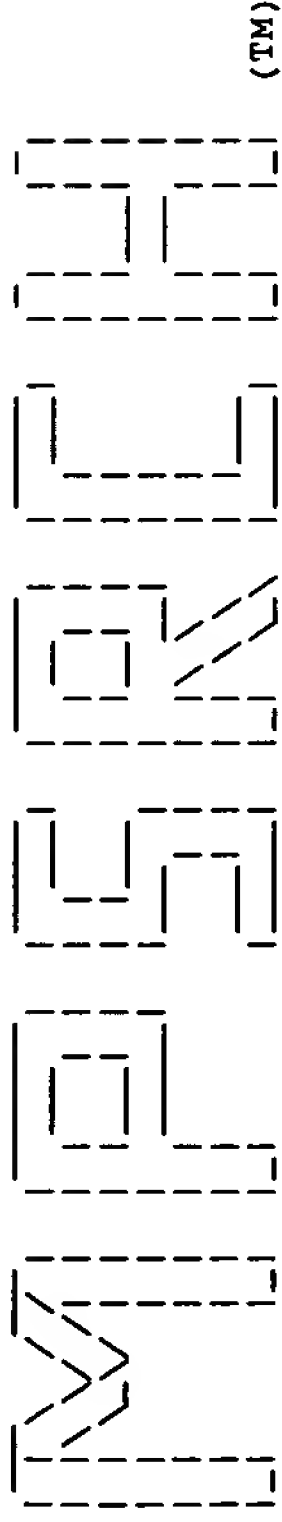
Query Match 8.5%; Score 100; DB 16; Length 485;  
Best Local Similarity 40.0%; Pred. No. 5.89e+00;  
Matches 14; Conservative 8; Mismatches 10; Indels 3; Gaps 3;

Db 43 wi-ppawkgtsqndvgygaydly-dlgefnqgkv 75  
I: | | | | : : | | | | : : | | | | : | | : | |  
QY 47 WVLGPVW-GTLYSAMGYGSYLVWKELGGFTEKAVV 80

Search completed: Tue Jul 21 14:53:06 1998  
Job time : 10 secs.



\*\*\*\*\*



\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

```

MPsrch_pp  protein - protein database search, using Smith-Waterman algorithm
Run on:      Tue Jul 21 14:53:23 1998;  MasPar time 10.79 Seconds
Tabular output not generated.  572.052 Million cell updates/sec

```

```

Title:
Description:
Perfect Score:
Sequence:

```

Scoring table: PAM 150  
Gap 11

Searched: 120441 seqs, 36531193 residues

Post-processing: Minimum Match 08  
Listing first 45 summaries

```
Database:      plr56
               1:plr1 2:plr2 3:plr3 4:plr4 5:nrl3d
```

**Statistics:** Mean 41.783; Variance 103.237; scale 0.405

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	1160	99.1	169	2	S14257	benzodiazepine recept	3.12e-170
2	1008	86.2	169	2	I57953	peripheral-type benzo	1.75e-144
3	1000	85.5	169	2	JC1393	benzodiazepine recept	3.93e-143
4	995	85.0	169	2	A53405	peripheral-type benzo	2.75e-142
5	956	81.7	169	2	I38105	peripheral benzodiaz	1.04e-135
6	944	80.7	169	2	A39473	peripheral-type benzo	1.10e-133
7	339	29.0	160	2	S04404	crtk protein - Rhodob	2.20e-34
8	327	27.9	153	2	B69434	benzodiazepine recept	1.60e-32
9	307	26.2	158	2	A57438	tryptophan-rich senso	1.94e-29
10	291	24.9	157	2	S49622	crtk protein - Rhodob	5.41e-27
11	211	18.0	31	2	I48082	mitochondrial benzodi	3.98e-15
12	160	13.7	155	2	G69987	hypothetical protein	4.83e-08
13	130	11.1	159	2	S76477	hypothetical protein	3.58e-04
14	115	9.8	639	2	S45776	uracil transport prot	2.33e-02
15	105	9.0	283	2	C47755	pectic enzyme secreti	3.28e-01
16	105	9.0	634	2	S75962	hypothetical protein	3.28e-01
17	102	8.7	188	2	S70680	bplJ protein - Bordet	7.08e-01
18	102	8.7	285	2	S77220	hypothetical protein	7.08e-01
19	101	8.6	246	2	D64800	glutamate/aspartate t	9.12e-01
20	100	8.5	292	2	G69897	conserved hypothetical	1.17e+00
21	100	8.5	368	2	A45831	MHC class I histocomp	1.17e+00
22	100	8.5	418	2	A64763	hypothetical protein	1.17e+00
23	97	8.3	271	2	JC4832	phosphatidate cytidyl	2.48e+00

24	97	8.3	518	1	A27705	alpha-amylase (EC 3.2	2.48e+00
25	95	8.1	637	2	F69869	heavy metal-transport	4.05e+00
26	94	8.0	759	2	S25330	SCR1 protein - yeast	5.16e+00
27	92	7.9	213	2	I51413	ORF - African clawed	8.35e+00
28	92	7.9	263	2	C39741	hypothetical 29K prot	8.35e+00
29	93	7.9	453	2	JQ2339	omega-3 fatty acid de	6.57e+00
30	92	7.9	677	2	S15179	choline transport sys	8.35e+00
31	92	7.9	700	2	E69146	sensory transduction	8.35e+00
32	91	7.8	245	2	E64989	heme exporter protein	1.06e+01
33	91	7.8	438	2	S40448	fumarate hydratase (E	1.06e+01
34	91	7.8	446	2	JQ2336	omega-3 fatty acid de	1.06e+01
35	91	7.8	533	2	S74401	hypothetical protein	1.06e+01
36	90	7.7	226	2	JQ2115	surface antigen - hep	1.34e+01
37	90	7.7	332	2	S77370	hypothetical protein	1.34e+01
38	90	7.7	389	2	S20749	surface antigen - hep	1.34e+01
39	90	7.7	434	2	S55155	probable tyrosine pho	1.34e+01
40	90	7.7	478	2	D64895	hypothetical protein	1.34e+01
41	89	7.6	226	2	JQ2077	surface antigen - hep	1.70e+01
42	89	7.6	308	2	S22923	ubiquinol--cytochrome	1.70e+01
43	89	7.6	378	2	S41870	surface antigen - hep	1.70e+01
44	89	7.6	460	2	G64916	hypothetical protein	1.70e+01
45	89	7.6	517	2	S21042	cytochrome-c oxidase	1.70e+01

## ALIGNMENTS

```

RESULT      1
ENTRY
TITLE      S14257 #type complete
ORGANISM   benzodiazepine receptor, peripheral-type - human
#formal_name Homo sapiens #common_name man
DATE       18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change
           20-Mar-1998
ACCESSIONS S14257
REFERENCE   S14257
#authors   Rlond, J.; Mattel, M.G.; Kaghad, M.; Dumont, X.; Guillemot
           J.C.; le Fur, G.; Caput, D.; Ferrara, P.
#journal   Eur. J. Biochem. (1991) 195:305-311
#title     Molecular cloning and chromosomal localization of a human
           peripheral-type benzodiazepine receptor.
#cross-references MUID:91146565
#accession S14257
           ##status      preliminary
           ##molecule_type mRNA
           ##residues    1-169 #label RIO
           ##cross-references GB:M36035; NID:gl84333; PID:g306883
KEYWORDS   mitochondrion; transmembrane protein
#length 169 #molecular-weight 18779 #checksum 4262
SUMMARY

```

	Query Match	99.1%;	Score 1160;	DB 2;	Length 169;
	Best Local Similarity	98.6%;	Pred. No. 3.12e-170;		
	Matches 141;	Conservative 2;	Mismatches 0;	Indels 0;	Gaps 0;
Db	27	HGEGLRWYAGLQKPSWHPHVVLPVVGTLYSAMGCGSYLVWKELGGFTEKAVVPLGLYT	86		
QY	27	HGEGLRWYAGLQKPSWHPHVVLPVVGTLYSAMGCGSYLVWKELGGFTEKAVVPLGLYT	86		
Db	87	GQALNWNWAPPIFFGARQMGWALVDLLLVSAAAATTVAWYQVSPLAARLLYPYLAWLAF	146		
QY	87	GQALNWNWAPPIFFGARQMGWALVDLLLVSAAAATTVAWYQVSPLAARLLYPYLAWLAF	146		
Db	147	ATTLNVCVWRDNHGWHGRRRLPE	169		
		:	:		
QY	147	TTTLNVCVWRDNHGWHGRRRLPE	169		

RESULT	2
ENTRY	
TITLE	157953 #type complete
ORGANISM	peripheral-type benzodiazepine receptor - mouse
DATE	#formal_name Mus musculus #common_name house mouse 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 02-Aug-1996
ACCESSIONS	157953
REFERENCE	157953

#authors Garnier, M.; Dimchev, A.B.; Boujrad, N.; Price, J.M.; Musto, N.A.; Papadopoulos, V.  
#journal Mol. Pharmacol. (1994) 45:201-211  
#title In vitro reconstitution of a functional peripheral-type benzodiazepine receptor from mouse Leydig tumor cells.  
#cross-references MUID:94158796  
#accession I57953  
##status preliminary; translated from GB/EMBL/DBJ  
##molecule\_type mRNA  
##residues 1-169 #label RES  
##cross-references GB:L17306; NID:g309441; PID:g309442  
#length 169 #molecular-weight 18841 #checksum 4647  
SUMMARY  
Query Match 86.2%; Score 1008; DB 2; Length 169;  
Best Local Similarity 84.6%; Pred. No. 1.75e-144;  
Matches 121; Conservative 12; Mismatches 10; Indels 0; Gaps 0;  
Db 27 RGEGLRWYASLQKPSWHPRTWLAPIGWTLYSAMGYSYVWKELGGFTEDAMVPLGLYT 86  
:|||||:|||||:|:|:|||||:|||||:|||||:|:|:|||||:|||||:|:|:|||||  
QY 27 HGEGLRWYAGLQKPSWHPHVPWGLVPTLYSAMGYSYLVWKELGGFTEKAVVPLGLYT 86  
Db 87 GOLALNWAWPPIFFGARQMGWALADLLLVSGVATATTIAWHRVSPPAARLLYPYLAWLAF 146  
|||||:|||||:|||||:|||||:|||||:|:|:|||||:|||||:|:|:|||||  
QY 87 GOLALNWAWPPIFFGARQMGWALVDLLLVSGAAATTVAWYQVSPLAARLLYPYLAWLAF 146  
Db 147 ATVLNYYVRDNGSGRRGSRLE 169  
:| ||| |||| |:| ||| ||||  
QY 147 TTTLNVCVRDNGHWRGRRRLPE 169  
RESULT 3 JC1393 #type complete  
ENTRY benzodiazepine receptor, mitochondrial - rat  
TITLE peripheral-type benzodiazepine receptor PKBS  
ALTERNATE\_NAMES #formal\_name Rattus norvegicus #common\_name Norway rat  
ORGANISM 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change  
DATE 10-Sep-1997  
ACCESSIONS JC1393; A32680  
REFERENCE JC1393  
#authors Casalotti, S.O.; Pelaia, G.; Yakovlev, A.G.; Csikos, T.; Grayson, D.R.; Krueger, K.E.  
#journal Gene (1992) 121:377-382  
#title Structure of the rat gene encoding the mitochondrial benzodiazepine receptor.  
#accession JC1393  
##molecule\_type DNA  
##residues 1-169 #label CAS  
##cross-references GB:M84221; NID:g206478; PID:g206480; GB:M95864  
REFERENCE A32680  
#authors Sprengel, R.; Werner, P.; Seeburg, P.H.; Mukhin, A.G.; Santi, M.R.; Grayson, D.R.; Guidotti, A.; Krueger, K.E.  
#journal J. Biol. Chem. (1989) 264:20415-20421  
#title Molecular cloning and expression of cDNA encoding a peripheral-type benzodiazepine receptor.  
#cross-references MUID:90062173  
#accession A32680  
##molecule\_type mRNA  
##residues 1-169 #label SPR  
##cross-references GB:J05122; NID:g206161; PID:g206162  
COMMENT This protein is located mainly in the mitochondrial outer membrane.  
GENETICS  
#gene MBR  
#introns 61/2; 107/3  
KEYWORDS mitochondrial; receptor; transmembrane protein  
FEATURE  
6-26 #domain transmembrane #status predicted #label TM1\  
47-67 #domain transmembrane #status predicted #label TM2\  
80-100 #domain transmembrane #status predicted #label TM3\  
106-126 #domain transmembrane #status predicted #label TM4\  
135-155 #domain transmembrane #status predicted #label TM5  
SUMMARY #length 169 #molecular-weight 18940 #checksum 5649  
Query Match 85.5%; Score 1000; DB 2; Length 169;

Best Local Similarity 83.2%; Pred. No. 3.93e-143;  
Matches 119; Conservative 14; Mismatches 10; Indels 0; Gaps 0;  
Db 27 RGEGLRWYASLQKPSWHPRTWLAPIGWTLYSAMGYSYIWKELGGFTEAMVPLGLYT 86  
:|||||:|||||:|:|:|||||:|||||:|||||:|:|:|||||:|||||:|:|:|||||  
QY 27 HGEGLRWYAGLQKPSWHPHVPWGLVPTLYSAMGYSYLVWKELGGFTEKAVVPLGLYT 86  
Db 87 GOLALNWAWPPIFFGARQMGWALVDMLVSGVATATTIAWHRVSPPAARLLYPYLAWLAF 146  
|||||:|||||:|||||:|||||:|||||:|:|:|||||:|||||:|:|:|||||  
QY 87 GOLALNWAWPPIFFGARQMGWALVDLLLVSGAAATTVAWYQVSPLAARLLYPYLAWLAF 146  
Db 147 ATMLNYYVRDNGSGRRGSRLE 169  
:| ||| |||| |:| ||| ||||  
QY 147 TTTLNVCVRDNGHWRGRRRLPE 169  
RESULT 4 A53405 #type complete  
ENTRY peripheral-type benzodiazepine receptor 1  
TITLE isoquinoline-binding protein - mouse  
ORGANISM #formal\_name Mus musculus #common\_name house mouse  
DATE 12-May-1994 #sequence\_revision 12-May-1994 #text\_change  
10-Sep-1997  
ACCESSIONS A53405; I55415; I49099  
REFERENCE A53405  
#authors Taketani, S.; Kohno, H.; Okuda, M.; Furukawa, T.; Tokunaga, R.  
#journal J. Biol. Chem. (1994) 269:7527-7531  
#title Induction of peripheral-type benzodiazepine receptors during differentiation of mouse erythroleukemia cells. A possible involvement of these receptors in heme biosynthesis.  
#accession A53405  
##status preliminary; not compared with conceptual translation  
##molecule\_type mRNA  
##residues 1-169 #label TAK  
##cross-references GB:D21207; NID:g484053; PID:d1005281; PID:g484054  
REFERENCE I55415  
#authors Taketani, S.; Kohno, H.; Okuda, M.; Furukawa, T.; Tokunaga, R.  
#journal J. Biol. Chem. (1994) 269:7527-7531  
#title Induction of peripheral-type benzodiazepine receptors during differentiation of mouse erythroleukemia cells. A possible involvement of these receptors in heme biosynthesis.  
#accession I55415  
##status preliminary; translated from GB/EMBL/DBJ  
##molecule\_type mRNA  
##residues 1-169 #label RES  
##cross-references GB:D21207; NID:g484053; PID:g484054  
REFERENCE I38724  
#authors Yakovlev, A.G.; Ruffo, M.; Jurka, J.; Krueger, K.E.  
#journal Gene (1995) 155:201-205  
#title Comparison of repetitive elements in the third intron of human and rodent mitochondrial benzodiazepine receptor-encoding genes.  
#cross-references MUID:95237610  
#accession I49099  
##status preliminary; translated from GB/EMBL/DBJ  
##molecule\_type DNA  
##residues 95-125 #label RE2  
##cross-references EMBL:U12419; NID:g529941; PID:g529942  
GENETICS  
#gene MBR  
#introns 107/3  
KEYWORDS mitochondrion  
SUMMARY #length 169 #molecular-weight 18829 #checksum 3513  
Query Match 85.0%; Score 995; DB 2; Length 169;  
Best Local Similarity 83.2%; Pred. No. 2.75e-142;  
Matches 119; Conservative 14; Mismatches 10; Indels 0; Gaps 0;  
Db 27 RGEGLRWYASLQKPSWHPRTWLAPIGWTLYSAMGYSYIWKELGGFTEDAMVPLGLYT 86  
:|||||:|||||:|:|:|||||:|||||:|||||:|:|:|||||:|||||:|:|:|||||  
QY 27 HGEGLRWYAGLQKPSWHPHVPWGLVPTLYSAMGYSYLVWKELGGFTEKAVVPLGLYT 86











```
##residues 1-639 ##label DEW
##cross-references EMBL:X78214; NID:g463261; PID:g872301
##experimental_source strain S288C
GENETICS
#map_position 2L
KEYWORDS  transmembrane protein
FEATURE
162-181      #domain transmembrane #status predicted #label TM1\
199-218      #domain transmembrane #status predicted #label TM2\
222-238      #domain transmembrane #status predicted #label TM3\
257-278      #domain transmembrane #status predicted #label TM4\
289-305      #domain transmembrane #status predicted #label TM5\
332-348      #domain transmembrane #status predicted #label TM6\
364-393      #domain transmembrane #status predicted #label TM7\
417-433      #domain transmembrane #status predicted #label TM8\
460-478      #domain transmembrane #status predicted #label TM9\
482-503      #domain transmembrane #status predicted #label TM10\
537-554      #domain transmembrane #status predicted #label TM11\
570-590      #domain transmembrane #status predicted #label TM12\
SUMMARY      #length 639 #molecular-weight 72164 #checksum 158

Query Match      9.8%; Score 115; DB 2; Length 639;
Best Local Similarity 26.2%; Pred. No. 2.33e-02;
Matches 22; Conservative 26; Mismatches 26; Indels 10; Gaps 8;

Db 151 TGLQLGLNW-WQT-WICIW-VGYTFVAFLLIGSKVGNN--YHISFPISRSVSFGIYFS 204
QY 86 TG-QLALNWAWPPIFFGARQMGWALVDLLVSGAAATTVAWYQVS-PLAARLLYP-YLA 142

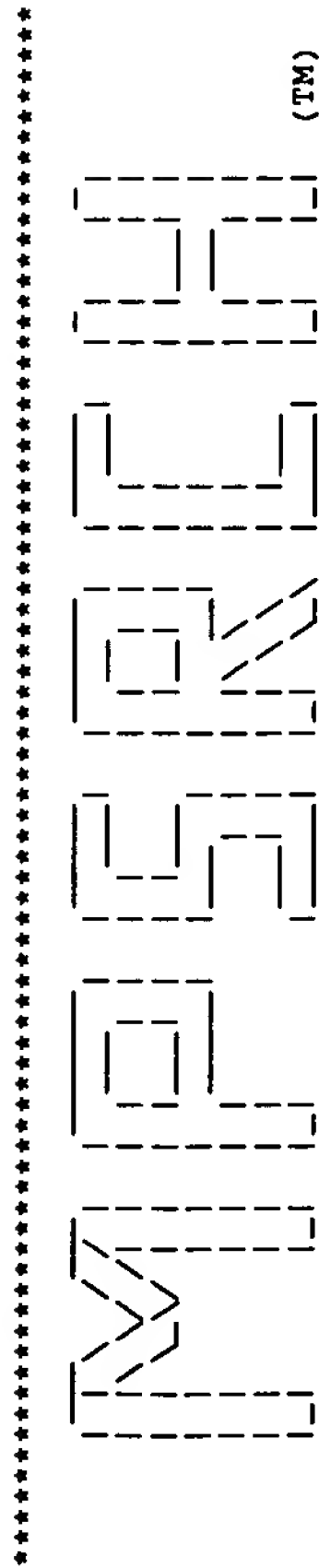
Db 205 IWIVINRVVMACVWNSTLAYIGSQ 228
QY 143 -WLAFTTTLNYCVWRDNRHGWRRGR 165

RESULT 15
ENTRY      C47755      #type complete
TITLE      pectic enzyme secretion protein Outo - Erwinia chrysanthemi
ORGANISM   #formal_name Erwinia chrysanthemi
DATE       24-Feb-1994 #sequence_revision 18-Nov-1994 #text_change
          21-Nov-1997
ACCESSIONS C47755
REFERENCE  A47021
#authors   Lindeberg, M.; Collmer, A.
#journal   J. Bacteriol. (1992) 174:7385-7397
#title     Analysis of eight out genes in a cluster required for pectic
          enzyme secretion by Erwinia chrysanthemi: sequence
          comparison with secretion genes from other gram-negative
          bacteria.
#cross-references MUID:93054355
#accession  C47755
#status     preliminary; not compared with conceptual translation
#molecule_type DNA
#residues   1-283 ##label LIN
#note       sequence extracted from NCBI backbone (NCBIP:118281)
CLASSIFICATION #superfamily hypothetical protein b2972
KEYWORDS      transmembrane protein
SUMMARY       #length 283 #molecular-weight 31355 #checksum 9186

Query Match      9.0%; Score 105; DB 2; Length 283;
Best Local Similarity 26.4%; Pred. No. 3.28e-01;
Matches 14; Conservative 23; Mismatches 13; Indels 3; Gaps 3;

Db 220 LGAW-LGWQALPNLVLIASLTGLTATLLWQRIHRLSMQOPLAFGPWLAVSGAM 271
QY 100 FGARQMGW-ALVDLLVSGAA-ATTVAWYQVSPLAARLLYPYLAWLAFTTL 150

Search completed: Tue Jul 21 14:53:50 1998
Job time : 27 secs.
```



Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Jul 21 14:54:08 1998; Maspar time 6.91 Seconds  
Tabular output not generated. 613.189 Million cell updates/sec

Title: >US-09-047-652A-3  
Description: (1-169) from US09047652A.pep  
Perfect Score: 1170  
Sequence: 1 XXXXXXXXXXXXXXXXXXXX.....LNYCVWRDNHGWRGRRRLPE 169

Scoring table: PAM 150  
Gap 11  
Searched: 69111 seqs, 25083644 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot35  
1:swiss1

Statistics: Mean 43.233; Variance 89.342; scale 0.484

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match	Length	DB ID	Description	Pred. No.
1	1160	99.1	169	1	PKBS_HUMAN PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR (PBR) (PKBS) (MITOCHONDRIAL)	3.59e-203
2	1008	86.2	169	1	PKBS_MOUSE PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR (PBR) (PKBS) (MITOCHONDRIAL)	1.94e-172
3	1000	85.5	169	1	PKBS_RAT PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR (PBR) (PKBS) (MITOCHONDRIAL)	7.93e-171
4	944	80.7	169	1	PKBS_BOVIN PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR (PBR) (PKBS) (MITOCHONDRIAL)	1.48e-159
5	339	29.0	160	1	CRTK_RHOCA CRK PROTEIN	3.11e-41
6	115	9.8	639	1	YBE2_YEAST PUTATIVE TRANSPORTER Y	2.22e-03
7	105	9.0	283	1	LEP3_ERWCH TYPE 4 PREPILIN-LIKE P	4.81e-02
8	101	8.6	246	1	GLTJ_ECOLI GLUTAMATE/ASPARTATE TR	1.58e-01
9	100	8.5	418	1	MHPT_ECOLI PUTATIVE 3-HYDROXYPHEN	2.11e-01
10	97	8.3	271	1	CDSA_PSEAE PHOSPHATIDATE CYTIDYLY	5.01e-01
11	97	8.3	518	1	AMT6_BACS7 GLUCAN 1,4-ALPHA-MALTO	5.01e-01
12	95	8.1	447	1	FD3C_SESIN OMEGA-3 FATTY ACID DES	8.83e-01
13	95	8.1	589	1	FUR4_SCHPO URACIL PERMEASE	8.83e-01
14	94	8.0	759	1	SCT1_YEAST CTR1 SUPPRESSOR PROTEI	1.17e+00
15	92	7.9	213	1	IF4E_XENLA EUKARYOTIC TRANSLATION	2.03e+00
16	92	7.9	263	1	CCMC_BRAJA HEME EXPORTER PROTEIN	2.03e+00
17	93	7.9	453	1	FD3C_SOYBN OMEGA-3 FATTY ACID DES	1.54e+00
18	92	7.9	677	1	BETT_ECOLI HIGH-AFFINITY CHOLINE	2.03e+00
19	91	7.8	245	1	CCMC_ECOLI HEME EXPORTER PROTEIN	2.68e+00
20	91	7.8	438	1	FD3C_SULSO FUMARATE HYDRATASE CLA	2.68e+00
21	91	7.8	446	1	FD3C_ARATH OMEGA-3 FATTY ACID DES	2.68e+00
22	91	7.8	498	1	OUSL_ERWCH OSMOPROTECTANT UPTAKE	2.68e+00
23	90	7.7	434	1	YNM8_YEAST HYPOTHETICAL 50.2 KD P	3.51e+00

24	89	7.6	308	1	CYB_COLRU CYTOCHROME B (EC 1.10.4.60e+00
25	89	7.6	517	1	QOX1_SULAC QUINOL OXIDASE POLYPEP
26	88	7.5	318	1	LPLB_BACSU LPLB PROTEIN
27	88	7.5	381	1	OPS2_DROME OPSIN RH2 (OCCELLAR OPS
28	88	7.5	389	1	VMSA_HPBWV MAJOR SURFACE ANTIGEN
29	88	7.5	400	1	VMSA_HPBV9 MAJOR SURFACE ANTIGEN
30	88	7.5	400	1	VMSA_HPBV2 MAJOR SURFACE ANTIGEN
31	88	7.5	859	1	OBP_HSVBC REPLICATION ORIGIN BIN
32	88	7.5	2127	1	RRPL_RABVS RNA POLYMERASE BETA SU
33	88	7.5	2142	1	RRPL_RABVP RNA POLYMERASE BETA SU
34	87	7.4	196	1	VATL_DICDI VACUOLAR ATP SYNTHASE
35	87	7.4	225	1	US07_HCMVA HYPOTHETICAL PROTEIN H
36	87	7.4	226	1	VMSA_HPBVD MAJOR SURFACE ANTIGEN
37	87	7.4	226	1	VMSA_HPBVS MAJOR SURFACE ANTIGEN
38	87	7.4	400	1	VMSA_HPBVP MAJOR SURFACE ANTIGEN
39	87	7.4	404	1	FD3C_BRANA OMEGA-3 FATTY ACID DES
40	87	7.4	414	1	IDH1_VIBAL ISOCITRATE DEHYDROGENA
41	86	7.4	443	1	YZ06_MYCTU HYPOTHETICAL 43.6 KD P
42	86	7.4	457	1	NU4M_ARBLI NADH-UBIQUINONE OXIDOR
43	87	7.4	496	1	NANT_ECOLI PUTATIVE SIALIC ACID T
44	87	7.4	753	1	PMT3_YEAST DOLICHYL-PHOSPHATE-MAN
45	87	7.4	857	1	ENV_HV2KR ENVELOPE POLYPROTEIN G

ALIGNMENTS

RESULT 1  
ID PKBS\_HUMAN STANDARD; PRT; 169 AA.  
AC P30536;  
DT 01-APR-1993 (REL. 25, CREATED)  
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)  
DE PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR (PBR) (PKBS) (MITOCHONDRIAL)  
DE BENZODIAZEPINE RECEPTOR)  
GN BZRP OR MBR.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; PRIMATES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 91146565.  
RA RIOND J., MATTEI M.G., KAGHAD M., DUMONT X., GUILLEMOT J.C.,  
RA LE FUR G., CAPUT D., FERRARA P.;  
RL EUR. J. BIOCHEM. 195:305-311(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 95237610.  
RA YAKOVLEV A.G., RUFFO M., JURKA J., KRUEGER K.E.;  
RL GENE 155:201-205(1995).  
CC -1- FUNCTION: RESPONSIBLE FOR THE MANIFESTATION OF PERIPHERAL-TYPE  
CC BENZODIAZEPINE RECOGNITION SITES AND IS MOST LIKELY TO COMPRISE  
CC BINDING DOMAINS FOR BENZODIAZEPINES AND ISOQUINOLINE CARBOXAMIDES.  
CC MAY PLAY A ROLE IN THE TRANSPORT OF PORPHYRINS AND HEME.  
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIUM; INTEGRAL MEMBRANE PROTEIN.  
CC -1- SIMILARITY: SEGMENTS TM1, TM4, AND TM5 SHOW SIMILARITY WITH THE  
CC TRANSMEMBRANE SEGMENTS M1, M2, AND M4 FROM SUBUNITS OF THE  
CC GABA(A)/BENZODIAZEPINE RECEPTOR FAMILY, RESPECTIVELY.  
DR EMBL; M36035; G306883; -  
DR EMBL; U12421; G529946; -  
DR PIR; S14257; S14257.  
DR MIM; 109610; -  
KW MITOCHONDRIUM; RECEPTOR; TRANSMEMBRANE.  
FT TRANSMEM 6 26 TM1 (POTENTIAL).  
FT TRANSMEM 47 67 TM2 (POTENTIAL).  
FT TRANSMEM 80 100 TM3 (POTENTIAL).  
FT TRANSMEM 106 126 TM4 (POTENTIAL).  
FT TRANSMEM 135 155 TM5 (POTENTIAL).  
FT CONFLICT 147 147 A -> T (IN REF. 2).  
FT CONFLICT 162 162 H -> R (IN REF. 2).  
SQ SEQUENCE 169 AA; 18779 MW; EC07D06E CRC32;

Query Match 99.1%; score 1160; DB 1; Length 169;  
Best Local Similarity 98.6%; Pred. No. 3.59e-203;

Matches 141; Conservative 2; Mismatches 0; Indels 0; Gaps 0;	
Db	27 HGEGLRWYAGLQKPSWHPHVLGPVWGTLYSAMGYGSYLVWKELGGFTEKAVVPLGLYT 86 :     :     :     :     :     :     :     :     :     :
QY	27 HGEGLRWYAGLQKPSWHPHVLGPVWGTLYSAMGYGSYLVWKELGGFTEKAVVPLGLYT 86 :     :     :     :     :     :     :     :     :     :
Db	87 GOLALNWAWPPIFFGARQMGWALVDLLVSGAAATTVAWYQVSPPLAARLLYPYLAWLAF 146 :     :     :     :     :     :     :     :     :     :
QY	87 GOLALNWAWPPIFFGARQMGWALVDLLVSGAAATTVAWYQVSPPLAARLLYPYLAWLAF 146 :     :     :     :     :     :     :     :     :     :
Db	147 ATTLNVCVWRDNHGWGRRRLPE 169 :     :     :     :     :     :     :     :     :     :
QY	147 TTTLNVCVWRDNHGWGRRRLPE 169 :     :     :     :     :     :     :     :     :     :
RESULT 2	
ID	PKBS_MOUSE STANDARD; PRT; 169 AA.
AC	P50637;
DT	01-OCT-1996 (REL. 34, CREATED)
DT	01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DE	PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR (PBR) (PKBS) (MITOCHONDRIAL
DE	BENZODIAZEPINE RECEPTOR).
GN	BZRP OR MBR.
OS	MUS MUSCULUS (MOUSE).
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC	EUTHERIA; RODENTIA.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE-TESTIS;
RX	MEDLINE; 94158796.
RA	GARNIER M., DIMCHEV A.B., BOUJRAD N., PRICE J.M., MUSTO N.A.,
RA	PAPADOPOULOS V.;
RL	MOL. PHARMACOL. 45:201-211(1994).
CC	-1- FUNCTION: RESPONSIBLE FOR THE MANIFESTATION OF PERIPHERAL-TYPE
CC	BENZODIAZEPINE RECOGNITION SITES AND IS MOST LIKELY TO COMPRISE
CC	BINDING DOMAINS FOR BENZODIAZEPINES AND ISOQUINOLINE CARBOXAMIDES.
CC	MAY PLAY A ROLE IN THE TRANSPORT OF PORPHYRINS AND HEME.
CC	-1- SUBCELLULAR LOCATION: MITOCHONDRION; INTEGRAL MEMBRANE PROTEIN.
CC	-1- SIMILARITY: SEGMENTS TM1, TM4, AND TM5 SHOW SIMILARITY WITH THE
CC	TRANSMEMBRANE SEGMENTS M1, M2, AND M4 FROM SUBUNITS OF THE
CC	GABA(A)/BENZODIAZEPINE RECEPTOR FAMILY, RESPECTIVELY.
DR	EMBL; L17306; G309442; -.
DR	MGI:88222; BZRP.
KW	MITOCHONDRION; RECEPTOR; TRANSMEMBRANE.
FT	TRANSMEM 6 26 TM1 (POTENTIAL).
FT	TRANSMEM 47 67 TM2 (POTENTIAL).
FT	TRANSMEM 80 100 TM3 (POTENTIAL).
FT	TRANSMEM 106 126 TM4 (POTENTIAL).
FT	TRANSMEM 135 155 TM5 (POTENTIAL).
SQ	SEQUENCE 169 AA; 18841 MW; C642B263 CRC32;
Query Match 86.2%; Score 1008; DB 1; Length 169;	
Best Local Similarity 84.6%; Pred. No. 1.94e-172;	
Matches 121; Conservative 12; Mismatches 10; Indels 0; Gaps 0;	
Db	27 RGEGLRWYASLQKPSWHPRWTLPITWGTLYSAMGYGSYLVWKELGGFTEKAVVPLGLYT 86 :     :     :     :     :     :     :     :     :     :
QY	27 HGEGLRWYAGLQKPSWHPHVLGPVWGTLYSAMGYGSYLVWKELGGFTEKAVVPLGLYT 86 :     :     :     :     :     :     :     :     :     :
Db	87 GOLALNWAWPPIFFGARQMGWALVDLLVSGVATATTLAWHRVSPPLAARLLYPYLAWLAF 146 :     :     :     :     :     :     :     :     :     :
QY	87 GOLALNWAWPPIFFGARQMGWALVDLLVSGAAATTVAWYQVSPPLAARLLYPYLAWLAF 146 :     :     :     :     :     :     :     :     :     :
Db	147 ATVLNVCVWRDNHGWGRRRLPE 169 :     :     :     :     :     :     :     :     :     :
QY	147 TTTLNVCVWRDNHGWGRRRLPE 169 :     :     :     :     :     :     :     :     :     :
RESULT 3	
ID	PKBS_RAT STANDARD; PRT; 169 AA.
AC	P16257;
DT	01-AUG-1990 (REL. 15, CREATED)

DT	01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE	PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR (PBR) (PKBS) (MITOCHONDRIAL
DE	BENZODIAZEPINE RECEPTOR).
GN	BZRP OR MBR.
OS	RATTUS NORVEGICUS (RAT).
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC	EUTHERIA; RODENTIA.
RN	[1]
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC	TISSUE-ADRENAL GLAND;
RX	MEDLINE; 90062173.
RA	SPRENGEL R., WERNER P., SEEBURG P.H., MUKHIN A.G., SANTI M.R.,
RA	GRAYSON D.R., GUIDOTTI A., KRUEGER K.E.;
RL	J. BIOL. CHEM. 264:20415-20421(1989).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN-WISTAR;
RX	MEDLINE; 93077057.
RA	CASALOTTI S.O., PELATA G., YAKOVLEV A.G., CSIKOS T., GRAYSON D.R.,
RA	KRUEGER K.E.;
RL	GENE 121:377-382(1992).
CC	-1- FUNCTION: RESPONSIBLE FOR THE MANIFESTATION OF PERIPHERAL-TYPE
CC	BENZODIAZEPINE RECOGNITION SITES AND IS MOST LIKELY TO COMPRISE
CC	BINDING DOMAINS FOR BENZODIAZEPINES AND ISOQUINOLINE CARBOXAMIDES.
CC	MAY PLAY A ROLE IN THE TRANSPORT OF PORPHYRINS AND HEME.
CC	-1- SUBCELLULAR LOCATION: MITOCHONDRION; INTEGRAL MEMBRANE PROTEIN.
CC	-1- SIMILARITY: SEGMENTS TM1, TM4, AND TM5 SHOW SIMILARITY WITH THE
CC	TRANSMEMBRANE SEGMENTS M1, M2, AND M4 FROM SUBUNITS OF THE
CC	GABA(A)/BENZODIAZEPINE RECEPTOR FAMILY, RESPECTIVELY.
DR	EMBL; J05122; G206162; -.
DR	EMBL; M84221; G206480; -.
DR	PIR; A32680; A32680.
DR	PIR; JCI393; JCI393.
KW	MITOCHONDRION; RECEPTOR; TRANSMEMBRANE.
FT	MOD_RES 71 71 BLOCKED.
FT	TRANSMEM 6 26 TM1 (POTENTIAL).
FT	TRANSMEM 47 67 TM2 (POTENTIAL).
FT	TRANSMEM 80 100 TM3 (POTENTIAL).
FT	TRANSMEM 106 126 TM4 (POTENTIAL).
FT	TRANSMEM 135 155 TM5 (POTENTIAL).
SQ	SEQUENCE 169 AA; 18940 MW; 1C22C453 CRC32;
Query Match 85.5%; Score 1000; DB 1; Length 169;	
Best Local Similarity 83.2%; Pred. No. 7.93e-171;	
Matches 119; Conservative 14; Mismatches 10; Indels 0; Gaps 0;	
Db	27 RGEGLRWYASLQKPSWHPRWTLPITWGTLYSAMGYGSYIWKELGGFTEAMVPLGLYT 86 :     :     :     :     :     :     :     :     :     :
QY	27 HGEGLRWYAGLQKPSWHPHVLGPVWGTLYSAMGYGSYLVWKELGGFTEKAVVPLGLYT 86 :     :     :     :     :     :     :     :     :     :
Db	87 GOLALNWAWPPIFFGARQMGWALVDMLVSGVATATTLAWHRVSPPLAARLLYPYLAWLAF 146 :     :     :     :     :     :     :     :     :     :
QY	87 GOLALNWAWPPIFFGARQMGWALVDLLVSGAAATTVAWYQVSPPLAARLLYPYLAWLAF 146 :     :     :     :     :     :     :     :     :     :
Db	147 ATMLNYYVWRDMSGRRGSRLTE 169 :     :     :     :     :     :     :     :     :     :
QY	147 TTTLNVCVWRDNHGWGRRRLPE 169 :     :     :     :     :     :     :     :     :     :
RESULT 4	
ID	PKBS_BOVIN STANDARD; PRT; 169 AA.
AC	P30535;
DT	01-APR-1993 (REL. 25, CREATED)
DT	01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE	PERIPHERAL-TYPE BENZODIAZEPINE RECEPTOR (PBR) (PKBS) (ISOQUINOLINE-
DE	BINDING PROTEIN) (IBP).
GN	BZRP.
OS	BOS TAURUS (BOVINE).
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC	EUTHERIA; ARTIODACTYLA.
RN	[1]















KW INITIATION FACTOR; PROTEIN BIOSYNTHESIS; RNA-BINDING.  
SO SEQUENCE 213 AA; 24635 MW; 4F60C133 CRC32;

Query Match 7.9%; Score 92; DB 1; Length 213;  
Best Local Similarity 25.3%; Pred. No. 2.03e+00;  
Matches 24; Conservative 18; Mismatches 47; Indels 6; Gaps 6;

Db 20 TGOEIVSPDQYIKH-PLQNRWALWFEKNDKSKTWQ-ANLRISKFDTVEDFWALYNHHIQL 77  
QY 75 TEKAVVPLGLYTGQALNWAWPPIFF-GARQMGWALVDLLVSGAAAATTV-AWYQVSPL 132

Db 78 SSNLMSCDYSL-FKDGIE-PMWEDEKNKRGGRWL 110  
QY 133 AARLLYPYLAWLAFTTTLNVCVWRDNDHGWRGRRRL 167

Search completed: Tue Jul 21 14:54:26 1998  
Job time : 18 secs.

**THIS PAGE BLANK (UBPTO)**

\*\*\*\*\*  
MISRELA (TM)  
\*\*\*\*\*  
Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd  
\*\*\*\*\*  
MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Jul 21 14:54:44 1998; Maspar time 12.32 Seconds  
Tabular output not generated.  
Title: >US-09-047-652A-3  
Description: (1-169) from US09047652A.pep  
Perfect Score: 1170  
Sequence: 1 XXXXXXXXXXXXXXXXXXXX.....LNYCVWRDNGWGRGRLPE 169

Scoring table: PAM 150  
Gap 11  
Searched: 140542 seqs, 42109429 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: sprembl5  
1:sp\_fungi 2:sp\_human 3:sp\_invertebrate 4:sp\_mammal  
5:sp\_mhc 6:sp\_organelle 7:sp\_phage 8:sp\_plant  
9:sp\_bacteria 10:sp\_rodent 11:sp\_virus 12:sp\_vertebrate  
13:sp\_unclassified  
Statistics: Mean 42.037; Variance 94.668; scale 0.444  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	995	85.0	169	10	Q62118 BENZODIAZEPINE RECEPTOR	4.27e-156
2	956	81.7	169	2	Q13850 PERIPHERAL BENZODIAZEP	7.01e-149
3	327	27.9	153	9	O28797 MITOCHONDRIAL BENZODIA	8.96e-36
4	291	24.9	157	9	Q54791 CRT GENES	1.02e-29
5	245	20.9	553	3	Q93367 HYPOTHETICAL PROTEIN C	3.72e-22
6	211	18.0	31	10	Q60419 MITOCHONDRIAL BENZODIA	9.88e-17
7	160	13.7	155	9	O34694 YTAB PROTEIN	5.49e-09
8	130	11.1	159	9	P74502 HYPOTHETICAL 17.8 KD P	9.21e-05
9	105	9.0	634	9	Q55465 HYPOTHETICAL 69.4 KD P	1.55e-01
10	102	8.7	188	9	Q45384 PUTATIVE GTG START COD	3.58e-01
11	102	8.7	285	9	P73514 HYPOTHETICAL 32.0 KD P	3.58e-01
12	100	8.5	292	9	O34416 YOAV	6.21e-01
13	100	8.5	368	5	Q30838 MHC CLASS I PROTEIN PR	6.21e-01
14	99	8.5	380	8	O23802 PLASTID OMEGA-3 FATTY	8.16e-01
15	99	8.5	436	8	Q40118 DELTA-15 LINEOYL DESAT	8.16e-01
16	100	8.5	537	9	P71879 HYPOTHETICAL 57.1 KD P	6.21e-01
17	98	8.4	209	9	O06630 HYPOTHETICAL 21.6 KD P	1.07e+00
18	97	8.3	372	9	O05804 HYPOTHETICAL 40.5 KD P	1.40e+00
19	97	8.3	438	8	O04807 OMEGA-3 FATTY ACID DES	1.40e+00
20	96	8.2	160	8	Q39553 RNA POLYMERASE (FRAGME	1.83e+00

21	96	8.2	261	9	P94945	FWUG, FWUD AND FWUB GE	1.83e+00
22	96	8.2	354	3	O16444	C54F6.7 PROTEIN.	1.83e+00
23	95	8.1	321	9	P95137	HYPOTHETICAL 35.9 KD P	2.39e+00
24	95	8.1	398	8	O24626	FATTY ACID DESATURASE	2.39e+00
25	95	8.1	443	8	O23824	FATTY ACID DESATURASE	2.39e+00
26	95	8.1	637	9	O31688	YKVV PROTEIN.	2.39e+00
27	94	8.0	441	8	P93350	OMEGA-3 FATTY ACID DES	3.11e+00
28	94	8.0	581	1	O14035	HYPOTHETICAL 64.1 KD P	3.11e+00
29	93	7.9	535	9	Q53111	CBB3-TYPE CYTOCHROME O	4.04e+00
30	92	7.9	676	2	O15296	15S-LIPOXYGENASE.	5.24e+00
31	93	7.9	685	8	P93156	CELLULOSE SYNTHASE (FR	4.04e+00
32	92	7.9	700	9	O26460	SENSORY TRANSDUCTION H	5.24e+00
33	91	7.8	117	9	O32937	HYPOTHETICAL 12.4 KD P	6.78e+00
34	91	7.8	458	6	O21406	NADH DEHYDROGENASE SUB	6.78e+00
35	91	7.8	533	9	Q55179	HYPOTHETICAL 57.1 KD P	6.78e+00
36	90	7.7	332	9	P73433	HYPOTHETICAL 36.6 KD P	8.76e+00
37	90	7.7	389	11	Q67875	PRE S1/S ORF.	8.76e+00
38	90	7.7	432	9	P77540	BENZOATE MEMBRANE TRAN	8.76e+00
39	90	7.7	439	1	O13295	ALPHA-GALACTOSIDASE.	8.76e+00
40	90	7.7	478	9	P76103	FROM BASES 1496666 TO	8.76e+00
41	89	7.6	378	11	Q67955	HBV SURFACE PROTEINS (	1.13e+01
42	89	7.6	460	9	P77429	FROM BASES 1676082 TO	1.13e+01
43	89	7.6	461	3	P91197	SIMILAR TO LIGAND-GATE	1.13e+01
44	88	7.5	281	11	O41740	MIDDLE S PROTEIN.	1.45e+01
45	88	7.5	389	11	Q67890	HBV SURFACE PROTEINS.	1.45e+01

ALIGNMENTS

RESULT 1  
ID Q62118 PRELIMINARY; PRT; 169 AA.  
AC Q62118;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)  
DE BENZODIAZEPINE RECEPTOR, PERIPHERAL (PERIPHERAL-TYPE BENZODIAZEPINE  
DE RECEPTOR).  
GN BZRP.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; RODENTIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/C;  
RA TAKETANI S., KOHNO H., OKUDA M., FURUKAWA T., TOKUNAGA R.;  
RL J. BIOL. CHEM. 269:2527-7531(1994).  
DR EMBL; D21207; G484054; -  
DR MGD; MGI:88222; BZRP.  
SQ SEQUENCE 169 AA; 18829 MW; 90CAC703 CRC32;

Query Match	85.0%	Score	995;	DB	10;	Length	169;
Best Local Similarity	83.2%	Pred. No.	4.27e-156;				
Matches	119;	Conservative	14;	Mismatches	10;	Indels	0; Gaps 0;
Db	27	RGEGLRWYASLQKPSWHPRTWLAPIWATLYSANGYGSYIVWKELGGFTEDAMVPLGLYT	86				
QY	27	HGEGLRWYAGLQKPSWHPHVVLPVWGTLYSANGYGSYLVWKELGGFTERAIVPLGLYT	86				
Db	87	GQLALNWAWPPIFFGARQMGWALDLLLLVSGVATATTLAWHRVSPPAARLLPYLAWLAF	146				
QY	87	GQLALNWAWPPIFFGARQMGWALDLLLLVSGAAATTVAVYQVSPLAARLLPYLAWLAF	146				
Db	147	ATVLYVWRDNGSRGGRSLAE	169				
QY	147	TTLNYCVWRDNGWGRGRLPE	169				
RESULT 2							
ID Q13850							
AC Q13850;							
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)							
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)							
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)							





RA O'CALLAGHAN M., PARSONS J., PERCY C., RIFKEN L., ROOPRA A.,  
RA SAUNDERS D., SHOWNKEEN R., SMALDON N., SMITH A., SONNHAMMER E.,  
RA STADEN R., SULSTON J., THIERRY-MIEG J., THOMAS K., VAUDIN M.,  
RA VAUGHAN K., WATERSTON R., WATSON A., WEINSTOCK L.,  
RA WILKINSON-SPROAT J., WOHLDMAN P.;  
RL NATURE 368:32-38(1994).  
DR EMBL; 281048; E348255; -.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 553 AA; 61142 MW; A2548F6C CRC32;  
  
Query Match 20.9%; Score 245; DB 3; Length 553;  
Best Local Similarity 35.7%; Pred. No. 3.72e-22;  
Matches 46; Conservative 32; Mismatches 46; Indels 5; Gaps 5;  
  
Db 421 WAAALKKPNWPKDVRVYSAV-DLLTSLPLGYASYLVYKNGGGFDYNDTKIALGLYGASV 479  
QY 33 WYAGLQKPSWHPH-WVLGPVWGTL-YSAMGYSGYLVWKELGGFT-EKAVVPLGLYTGQL 89  
  
Db 480 TLAVATPIV-KKRELGLWKNTTVVSLTAAASFAFYKIDKKAGLLVVPFAVWTFAYAY 538  
QY 90 ALNWAWPPIFFGARQMGWALVDLLLVSGAAATVAVYQVSPLAARLLPYLAWLAFTTT 149  
  
Db 539 LAYSIKKEN 547  
QY 150 LNYCVWRDN 158  
  
RESULT 6  
ID Q60419 PRELIMINARY; PRT; 31 AA.  
AC Q60419;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)  
DE MITOCHONDRIAL BENZODIAZEPINE RECEPTOR (FRAGMENT).  
GN MBR.  
OS CRICETULUS GRISEUS (CHINESE HAMSTER).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; RODENTIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 95237610.  
RA YAKOVLEV A.G., RUFFO M., JURKA J., KRUEGER K.E.;  
RL GENE 155:201-205(1995).  
DR EMBL; U12420; G529944; -.  
FT NON\_TER 1 1  
FT NON\_TER 31 31  
SQ SEQUENCE 31 AA; 3275 MW; 3D86A9D4 CRC32;  
  
Query Match 18.0%; Score 211; DB 10; Length 31;  
Best Local Similarity 87.1%; Pred. No. 9.88e-17;  
Matches 27; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
Db 1 WPIIFFGARQMGWALADLLLVSGVATATTLA 31  
QY 95 WPIIFFGARQMGWALVDLLLVSGAAATTTVA 125  
  
RESULT 7  
ID O34694 PRELIMINARY; PRT; 155 AA.  
AC O34694;  
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)  
DE YTAB PROTEIN.  
GN YTAB.  
OS BACILLUS SUBTILIS.  
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,  
RA AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,  
RA BORRIS R., BOURSIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,

RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,  
RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,  
RA DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMMERSON P.T.,  
RA ENTIAN K.D., ERRINGTON J., FABRET C., FERRARI E., FOULGER D.,  
RA FRITZ C., FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N.,  
RA GHIM S.Y., GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G.,  
RA GUISEPPI G., GUY B.J., HAGA K., HAIECH J., HARWOOD C.R., HENAUT A.,  
RA HILBERT H., HOLSAPPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L.,  
RA JORIS B., KARAWATA D., KASAHARA Y., KLAERR-BLANCHARD M., KLEIN C.,  
RA KOBAYASHI Y., KOETTER P., KONINGSTEIN G., KROGH S., KUMANO M.,  
RA KURITA K., LAPIDUS A., LARDINOIS S., LAUBER J., LAZAREVIC V.,  
RA LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,  
RA MEDINA N., MELLADO R.P., MIZUNO M., MOESTL D., NAKAI S., NOBACK M.,  
RA NOONE D., O'REILLY M., OGAWA K., OGIWARA A., OUDEGA B., PARK S.H.,  
RA PARRO V., PUJIC P., PORTETELLE D., PORWOLLIK S., PRESCOTT A.M.,  
RA PRESECAN E., PUJIC P., PURNELLE B., RAPOPORT G., REY M., REYNOLDS S.,  
RA RIEGER M., RIVOLTA C., ROCHA E., ROCHE B., ROSE M., SADAIE Y.,  
RA SATO T., SCANLAN E., SCHLEICH S., SCHROETER R., SCOFFONE F.,  
RA SEKIGUCHI J., SEKOWSKA A., SEROR S.J., SERROR P., SHIN B.S., SOLDI B.,  
RA SOROKIN A., TACCONI E., TAKAGI T., TAKAHASHI H., TAKEMARU K.,  
RA TAKEUCHI M., TAMAKOSHI A., TANAKA T., TERPSTRA P., TOGNONI A.,  
RA TOSATO V., UCHIYAMA S., VANDENBOL M., VANNIER F., VASSAROTTI A.,  
RA VIARI A., WAMBUIT R., WEDLER E., WEDLER H., WEITZENEGGER T.,  
RA WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K., YASUMOTO K., YATA K.,  
RA YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E., YOSHIKAWA H., DANCHIN A.;  
RL NATURE 390:249-256(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;  
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBBJ DATA BANKS.  
RN [3]  
RP SEQUENCE OF 119-155 FROM N.A.  
RX MEDLINE; 94195107.  
RA KIEL J.A., BOELS J.M., BELDMAN G., VENEMA G.;  
RL MOL. MICROBIOL. 11:203-218(1994).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA LAPIDUS A., GALLERON N., SOROKIN A., EHRLICH D.;  
RL SUBMITTED (JUN-1997) TO EMBL/GENBANK/DBBJ DATA BANKS.  
DR EMBL; 299119; E1185966; -.  
DR EMBL; AF008220; G2293285; -.  
SQ SEQUENCE 155 AA; 17288 MW; 875DAD85 CRC32;  
  
Query Match 13.7%; Score 160; DB 9; Length 155;  
Best Local Similarity 27.9%; Pred. No. 5.49e-09;  
Matches 34; Conservative 29; Mismatches 57; Indels 2; Gaps 2;  
  
Db 31 WYNSLKKPDWTPSGTAIGIIWAILFALISLSAAIVYAAS-FKGAKEFWFTLLIN-YVLN 88  
QY 33 WYAGLQKPSWHPHWWVLGPVWGTLYSAMGYSGYLVWKELGGFTKAVVPLGLYTGQLALN 92  
  
Db 89 QAFSYFQFTQKNLLAASLDCLLVAITAVLIIIAKKYSRAASYLLLPFLWSAFATLSF 148  
QY 93 WAWPPIFFGARQMGWALVDLLLVSGAAATTVAVYQVSPLAARLLPYLAWLAFTTTTLY 152  
  
Db 149 TI 150  
QY 153 CV 154  
  
RESULT 8  
ID P74502 PRELIMINARY; PRT; 159 AA.  
AC P74502;  
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)  
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)  
DT 01-FEB-1997 (TREMBLREL. 02, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 17.8 KD PROTEIN.  
OS SYNECHOCYSTIS SP.  
OC EUBACTERIA; CYANOBACTERIA; CHROCOCCALES; SYNECHOCYSTIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PCC6803;

```

RA ALLEN A.G., MASKELL D.J.;
RL MOL. MICROBIOL. 19:37-52(1996).
DR EMBL; X90711; G992979; -
SQ SEQUENCE 188 AA; 21467 MW; F6C7644C CRC32;

Query Match      8.7%; Score 102; DB 9; Length 188;
Best Local Similarity 29.0%; Pred.No. 3.58e-01;
Matches 9; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

Db 121 TGGVWYNINDMATLVFANLIWLIIKKILSF 151
    | |:::|:::| |:::| |:::| |:::|
QY 122 TTVAWYQVSPLAARLLYPYLAWLAFTTLNY 152

RESULT 11
ID P73514 PRELIMINARY; PRT; 285 AA.
AC P73514;
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DE 01-FEB-1997 (TREMBLREL. 02, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 32.0 KD PROTEIN.
GN GLGP.
OS SYNECHOCYSTIS SP.
OC EUBACTERIA; CYANOBACTERIA; CHROCOCCALES; SYNECHOCYSTIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC6803;
RA TABATA S.;
RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC6803;
RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,
RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,
RA HOSEUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUO K.,
RA OKUMURA S., SHIMPO S., TAKEUCHI C., WADA T., WATANABE A.,
RA YAMADA M., YASUDA M., TABATA S.;
RL DNA RES. 3:109-136(1996).
DR EMBL; D90907; G1652634; -.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 285 AA; 31950 MW; BE7A9038 CRC32;

Query Match      8.7%; Score 102; DB 9; Length 285;
Best Local Similarity 30.6%; Pred.No. 3.58e-01;
Matches 26; Conservative 24; Mismatches 23; Indels 12; Gaps 11;

Db 137 AWVSPGQG-TTILWGAVFLAAG-G--LVLGKVGWRW-DTSIMFLAVYAGLDLARNALGW 191
    :|:|:| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 41 SWHPP-HWVLGPVW-GTLYSAMGYSLVWKELGGFTERAVVPLGLYTG-QIALN-W-AW 95

Db 192 PPEVSFHLENGSLLVFALEMLTDP 216
    ||:| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
QY 96 PP-IFFGAROMGWALV-DLLLVSGA 118

RESULT 12
ID O34416 PRELIMINARY; PRT; 292 AA.
AC O34416;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE YOAV.
GN YOAV.
OS BACILLUS SUBTILIS.
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RA LAPIDUS A., GALLERON N., SOROKIN A., EHRLICH D.;
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA KUNST F., OGASAWARA N., MOSTER T.
```



**THIS PAGE BLANK (USPTO)**



This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problems Mailbox.**

**THIS PAGE BLANK (USPTO)**